



TESIS SS14 - 2501

**MODEL *MIXTURE SURVIVAL* SPASIAL
PADA ANGKA LAMA SEKOLAH ANAK
UMUR 16-18 TAHUN
DI JAWA TIMUR TAHUN 2012**

**JIHADIL QUDSI S.
NRP 1313 201 033**

**DOSEN PEMBIMBING
Prof. Drs. Nur Iriawan, M.Ikom., Ph.D.**

**PROGRAM MAGISTER
JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2015**

(Halaman ini sengaja dikosongkan)



TESIS SS14 - 2501

MIXTURE SURVIVAL SPASIAL MODEL
ON CHILD EDUCATION TIME 16-18 YEARS OLD
IN EAST JAVA 2012

JIHADIL QUDSI S.
NRP 1313 201 033

DOSEN PEMBIMBING
Prof. Drs. Nur Iriawan, M.Ikom., Ph.D.

PROGRAM MAGISTER
JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2015

(Halaman ini sengaja dikosongkan)

**MODEL MIXTURE SURVIVAL SPASIAL
PADA ANGKA LAMA SEKOLAH ANAK UMUR 16-18 TAHUN
DI JAWA TIMUR TAHUN 2012**

**Tesis disusun untuk memenuhi salah satu syarat memperoleh gelar
Magister Sains (M.Si)**

di

Institut Teknologi Sepuluh Nopember

Oleh :


JIHADIL QUDSI S.

NRP. 1313201033

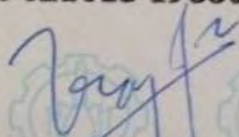
Tanggal Ujian : 5 Maret 2015

Periode Wisuda : September 2015

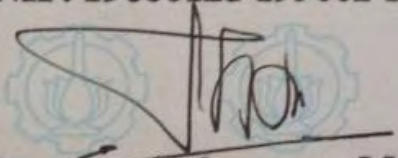
Disetujui Oleh :

1. 
Prof. Drs. Nur Iriawan, M.Ikom, Ph.D.
NIP. 19621015 198803 1 002

(Pembimbing)


2. 
Dr. Brodjol Sutijo Suprih Ulama, M.Si.
NIP. 19660125 199002 1 001

(Penguji)

3. 
Dr. Agus Suharsono, M.S.
NIP. 19580823 198403 1 003

(Penguji)

Direktur Pascasarjana ITS


Prof. Dr. Ir. Adi Soeprijanto, MT

NIP. 19640405 199002 1 001

MODEL *MIXTURE SURVIVAL* SPASIAL PADA ANGKA LAMA SEKOLAH ANAK UMUR 16-18 TAHUN DI JAWA TIMUR TAHUN 2012

Nama : Jihadil Qudsi S.
NRP : 1313 201 033
Pembimbing : Prof. Drs. Nur Iriawan, M.Ikom., Ph.D.

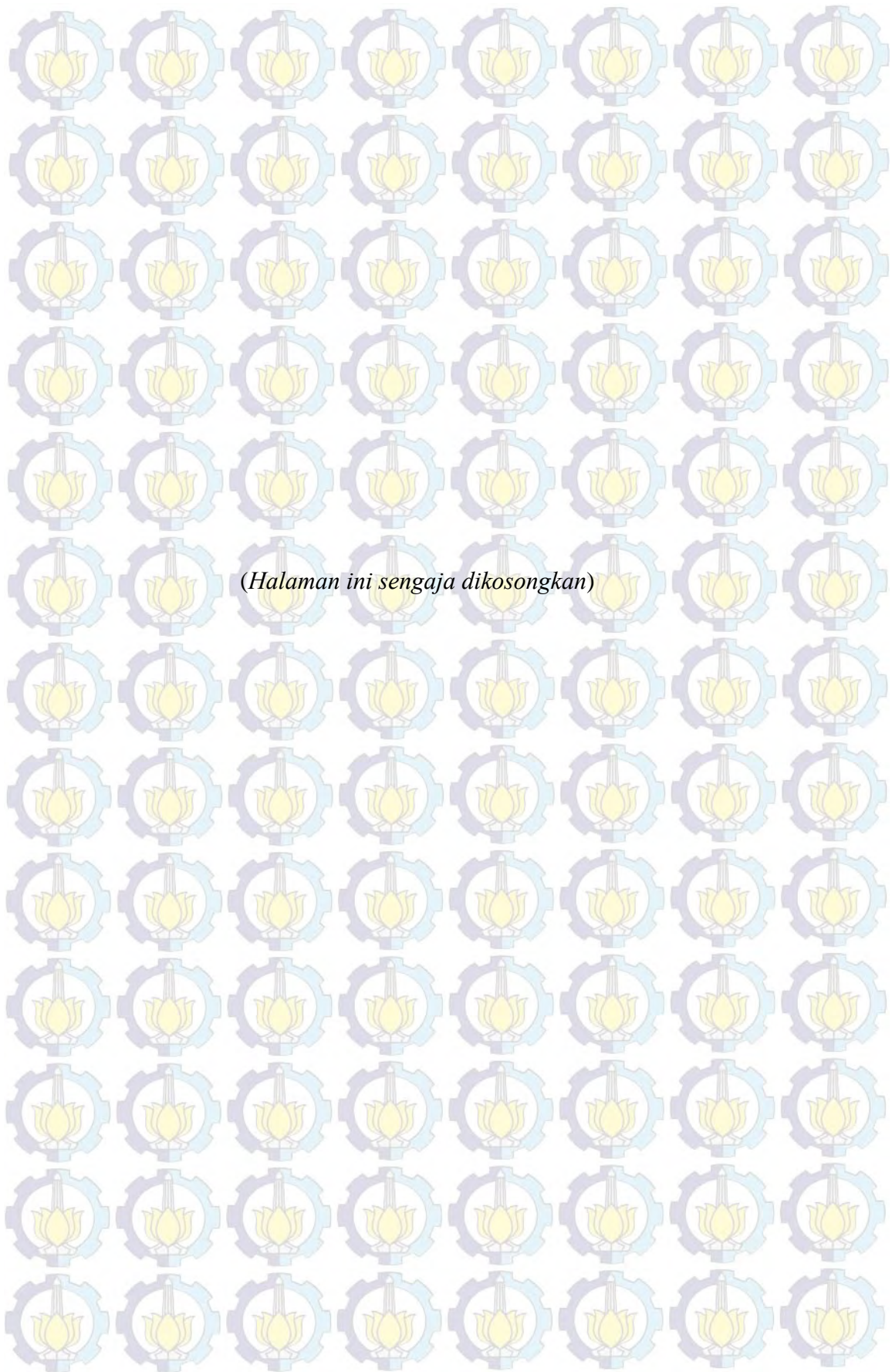
ABSTRAK

Analisis *survival* merupakan analisis mengenai data lama waktu suatu obyek sampai terjadinya peristiwa tertentu yang disebut sebagai *failure event*. Perkembangan analisis *survival* cukup pesat hingga diberikan penambahan efek spasial serta pemberian konsep *mixture* pada model *survival*. Penelitian ini bertujuan untuk memperoleh model *survival* dengan penambahan konsep *mixture* dan pertimbangan adanya efek spasial serta mencari variabel-variabel yang mempengaruhi model *mixture survival* spasial tersebut.

Data yang digunakan adalah data skunder dari Badan Pusat Statistik provinsi Jawa Timur, yang merupakan data hasil dari survei sosial ekonomi nasional (susenas) tahun 2012. Variabel respon yaitu angka lama sekolah anak umur 16-18 tahun se Jawa Timur yang menjadi responden pada susenas tahun 2012, sedangkan variabel prediktor untuk menyusun model yaitu jenis kelamin, status perkawinan, status bekerja, tingkat pendidikan kepala rumah tangga, jenis kelamin kepala rumah tangga, rata-rata pengeluaran rumah tangga per kapita, jumlah anggota rumah tangga, dan lokasi tempat tinggal (kota atau desa). Dalam penelitian ini akan ditentukan banyak komponen *mixture* dengan algoritma RJMCMC, setelah didapat banyak komponen, dilakukan estimasi parameter untuk mencari variabel-variabel yang secara signifikan mempengaruhi variabel respon dengan Gibbs Sampling menggunakan software WinBUGS.

Hasil RJMCMC didapat ada 6 komponen *mixture* yang paling optimum pada data angka lama sekolah dengan variabel-variabel prediktornya. Sedangkan hasil dari Gibbs Sampling didapat bahwa pada setiap komponen *mixture*, variabel prediktor yang secara signifikan mempengaruhi variabel respon berbeda-beda dan efek spasial dari daerah kabupaten/kota tidak semua secara signifikan mempengaruhi angka lama sekolah anak umur 16-18 tahun.

Kata Kunci: Analisis *Survival*, Angka Lama Sekolah, *Mixture Survival* Spasial RJMCMC.



MIXTURE SURVIVAL SPATIAL MODEL ON CHILD EDUCATION TIME 16-18 YEARS OLD IN EAST JAVA 2012

By : Jihadil Qudsi S.
Student Identity Number : 1313 201 033
Supervisor : Prof. Drs. Nur Iriawan, M.Ikom., Ph.D.

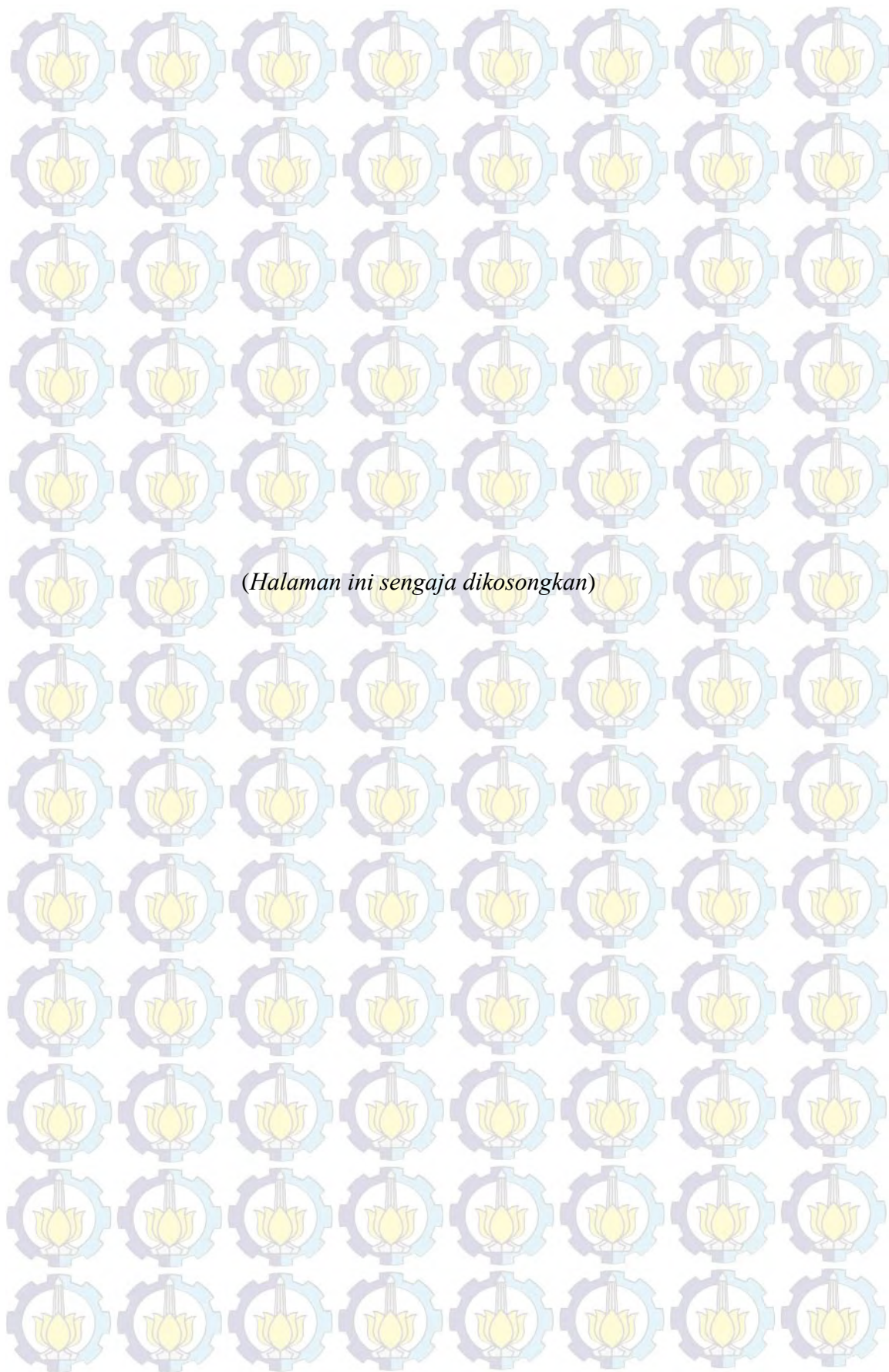
ABSTRACT

Survival analysis is a time data object to certain event that called failure event. This method has quickly development, also mixed with spatial effect and mixture concept given to applied in spatial case. This research aims to get survival model that accommodated mixture concept with spatial effect to determine the influence variable in the model.

This model applied on child education time 16-18 years old in East Java. The data came from SUSENAS 2012 by BPS. Variable that predicted influence child education time 16-18 years old are sex, marriage status, working status, head of household's education, head of household's sex, per capita income, number of household's members and residence location. The number of mixture components determine by RJMCMC algorithm and then parameter estimation to determine the influence factors significantly with Gibbs sampling by using WinBUGS.

Based on the optimum RJMCMC, there are 6 mixture components on child education time 16-18 years old with the predictor variables. While Gibbs Sampling resulted that each of mixture component has a significant different predictors that influence the response variable and the spatial effect from each region not at all influence child education time 16-18 years old significantly.

Key Words: Education Time, Survival Analysis, Mixture Survival Spatial, RJMCMC.





MOTTO DAN PERSEMBAHAN

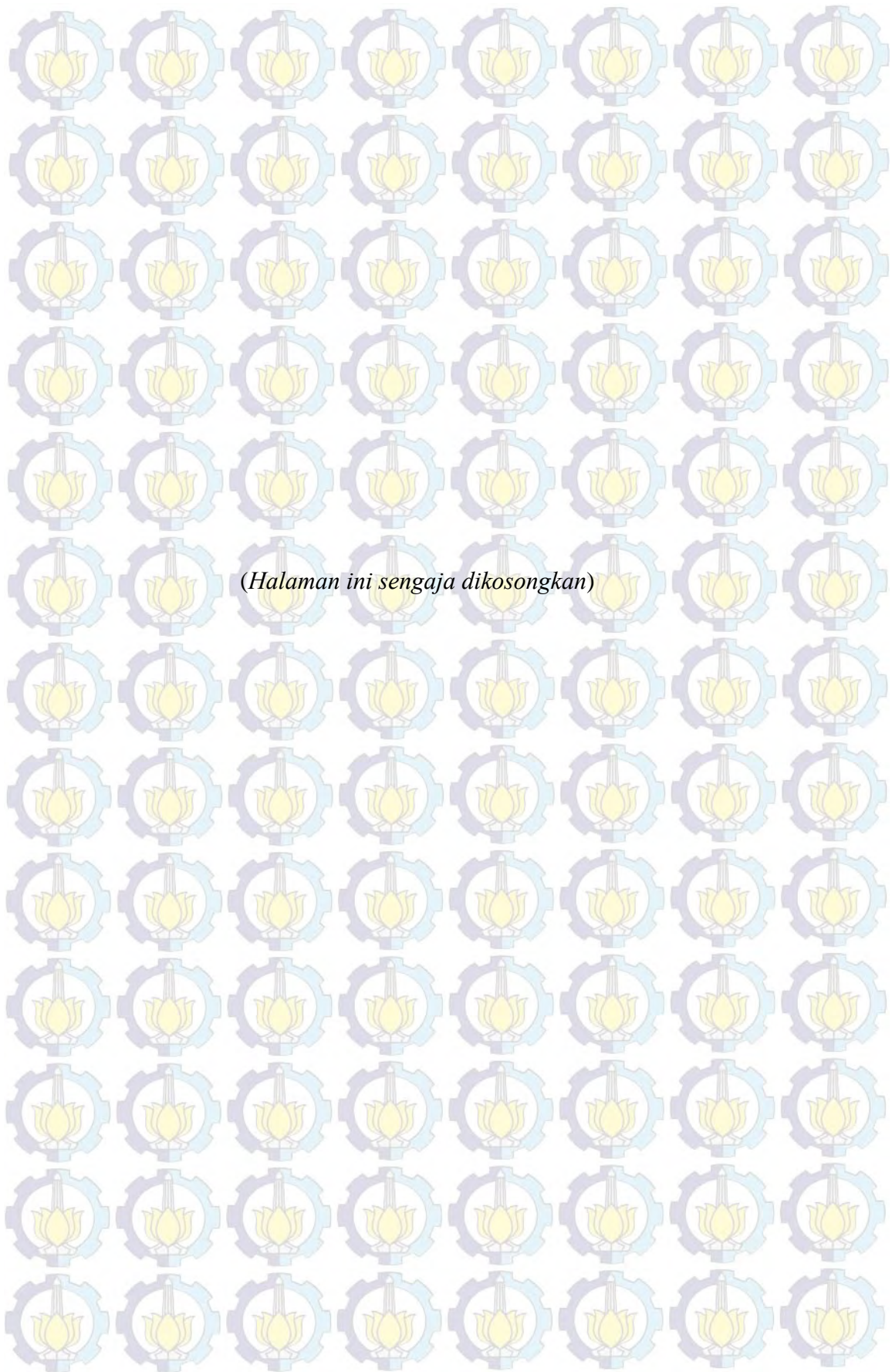
Allah SWT tidak akan memberikan cobaan yang tidak bisa dilalui oleh hambanya

Berbaik-sangka kepada Allah SWT
Dibalik cobaan ada hikmah yang terkandung

Karya tulis ini kupersembahkan

Untuk kedua orang tua ku tercinta
Bapak ku, Drs. Satarudin, M.Si., dan Mama ku, Nurhayati
Terimakasih atas perjuangan, air mata kasih dalam doa
Serta cinta dan kasih sayang tak terbalaskan selama ini.

Untuk adik-adikku tersayang
Firyal Dhiyaul Haqqi, Indah Rizqiana Amalia, dan Ardelia Mufarrihati
Terimakasih atas canda tawa, doa, selisih paham dan dukungannya



KATA PENGANTAR

Syukur Alhamdulillah kehadiran Allah SWT atas segala ridho, rahmat, nikmat, dan hidayah-Nya sehingga penulis dapat menyelesaikan Tesis dengan judul:

**MODEL *MIXTURE SURVIVAL* SPASIAL
PADA ANGKA LAMA SEKOLAH ANAK UMUR 16-18 TAHUN
DI JAWA TIMUR TAHUN 2012**

Sholawat serta salam senantiasa penulis curahkan kepada junjungan Nabi besar Muhammad SAW beserta para keluarga serta sahabatnya yang telah memberi petunjuk kepada seluruh umat manusia dalam menapaki jalan kehidupan sehingga selalu berada di jalan serta arah yang diridhoi oleh Allah SWT.

Selesaiya laporan Tesis ini tidak lepas dari bantuan, bimbingan serta dukungan dari berbagai pihak. Oleh karena itu penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Kedua orang tua tercinta, Bapak Drs. Satarudin, M.Si., dan Mama Nurhayati, atas kegigihan, pejuangan, air mata dalam doa, dan kasih sayanginya selama ini, dan juga kepada adik-adikku tersayang, Fery, Indah dan Ardelia, atas segenap cinta, dan dukungannya,
2. Bapak Prof. Drs. Nur Iriawan, M.Ikom Ph.D., selaku dosen pembimbing yang telah begitu banyak meluangkan waktu untuk memberikan bimbingan, arahan dan bantuan kepada penulis Tesis ini dapat terselesaikan,
3. Bapak Dr. Suhartono, M.Sc., selaku Ketua Program Studi Magister Statistika FMIPA ITS, yang telah memberikan dukungan kepada penulis untuk melaksanakan tesis ini,
4. Ibu Dra. Destri Susilaningrum, M.Si., selaku dosen wali yang telah memberikan bantuan kepada penulis selama ini,
5. Bapak Dr. Brojol Sutijo Suprih Ulama, M.Si dan Bapak Dr. Agus Suharsono, M.S. selaku dosen penguji yang telah memberikan saran dan arahan hingga terselesaikannya laporan Tesis ini,

6. Para dosen pengajar beserta segenap staf karyawan jurusan Statistika ITS yang telah memberikan ilmu-ilmu yang tiada ternilai harganya dan memberikan bantuan selama proses perkuliahan. Semoga jasa bapak dan ibu tercatat sebagai amal dan pahala.
7. Seorang wanita yang penulis sayangi, Fadhlina Apriyanti, S.Kep., yang selalu memberikan dukungan, dorongan, motivasi, dan semangat kepada penulis untuk terus berjuang hingga akhir,
8. Pak Rindang dan Pak Syahrul selaku mahasiswa S3 Statistika ITS yang telah memberikan banyak bantuan dan arahan kepada penulis,
9. Teman seperjuangan satu pembimbing, Kiki, Irma dan Mas Choiril yang telah berjuang bersama dalam menyelesaikan laporan Tesis,
10. Sahabat-sahabatku, Zul, Ina, Mei, Farida, Nariza, Kiki, Mas Amin dan Taurif yang telah bersedia untuk berbagi pengalaman dan memberikan dukungan serta mendengarkan keluhkesah penulis,
11. Seluruh mahasiswa S2 angkatan 2013, terimakasih atas pengalaman, kenangan dan dukungan kepada penulis,
12. Penghuni kos Keputih 3B/52, Mas Catur, Mas Erie, Afif, Mas Untung dan Mas Nora, atas pengalaman, kenangan dan suka cita,
13. Semua pihak yang telah membantu yang tak dapat disebutkan satu persatu.

Dengan selesainya laporan ini, penulis menyadari bahwa Tesis ini masih ada kekurangan, maka diharapkan saran dan kritik agar dapat mengembangkan penelitian Tesis ini. Semoga memberi manfaat kepada penulis, pembaca, dan penelitian selanjutnya. *Amin.*

Surabaya,

Maret 2015

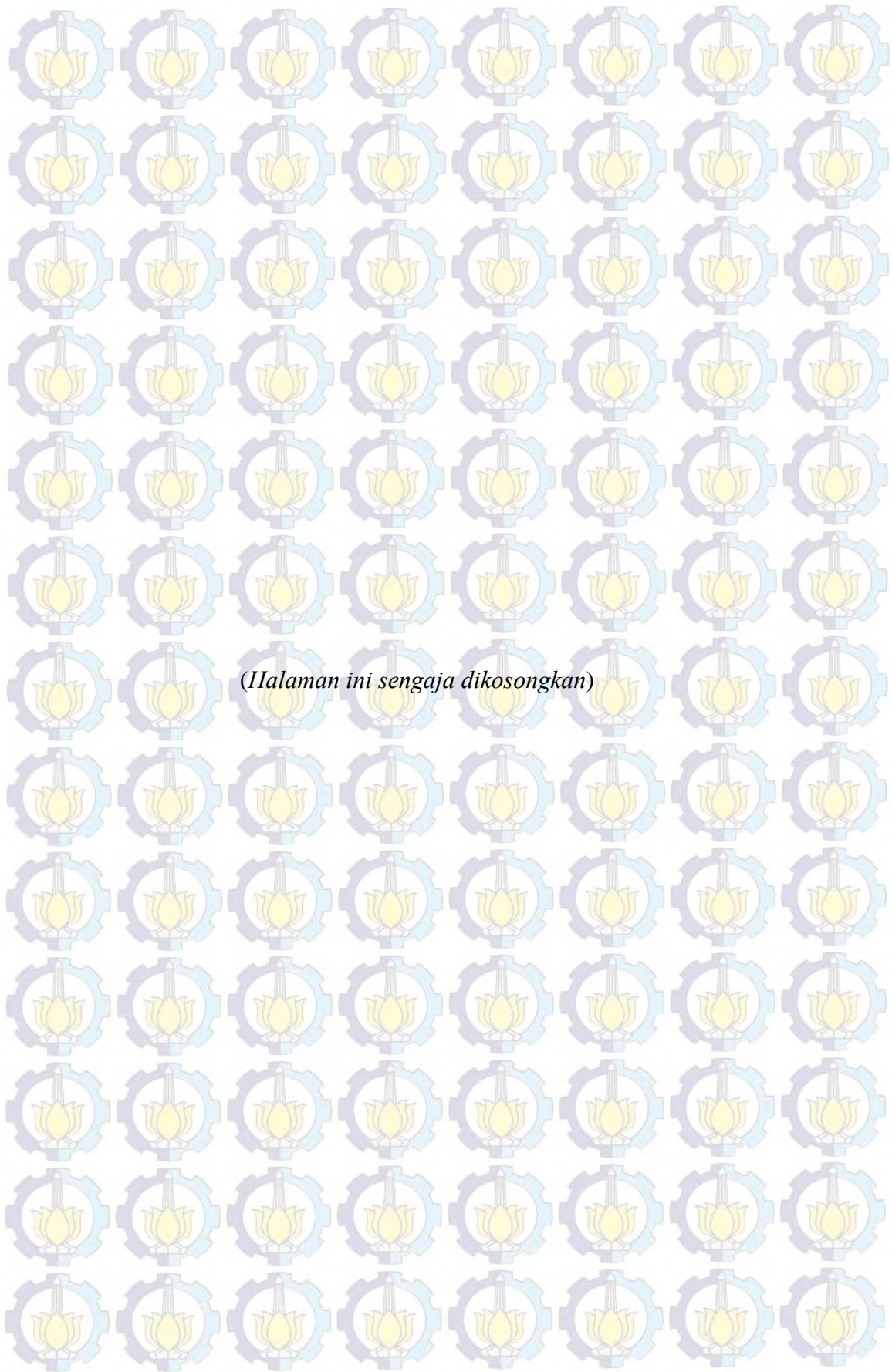
Penulis

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PENGESAHAN	iii
ABSTRAK	v
ABSTRACT	vi
MOTO DAN PERSEMBAHAN	ix
KATA PENGANTAR.....	xi
DAFTAR ISI.....	xiii
DAFTAR TABEL	xvii
DAFTAR GAMBAR.....	xix
DAFTAR LAMPIRAN	xxi
BAB 1 PENDAHULUAN	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah	5
1.3. Tujuan Penelitian.....	5
1.4. Manfaat Penelitian.....	5
1.5. Batasan Masalah	6
BAB II TINJAUAN PUSTAKA	7
2.1. Uji Goodness of Fit	7
2.2. Analisis Survival	8
2.2.1. Fungsi Survival dan Fungsi Hazard	10
2.2.2. Model dan Asumsi Hazard Proporsional.....	11
2.3. Model Survival Spasial.....	12
2.3.1. Model Geostatistik	14
2.3.2. Model Lattice	14

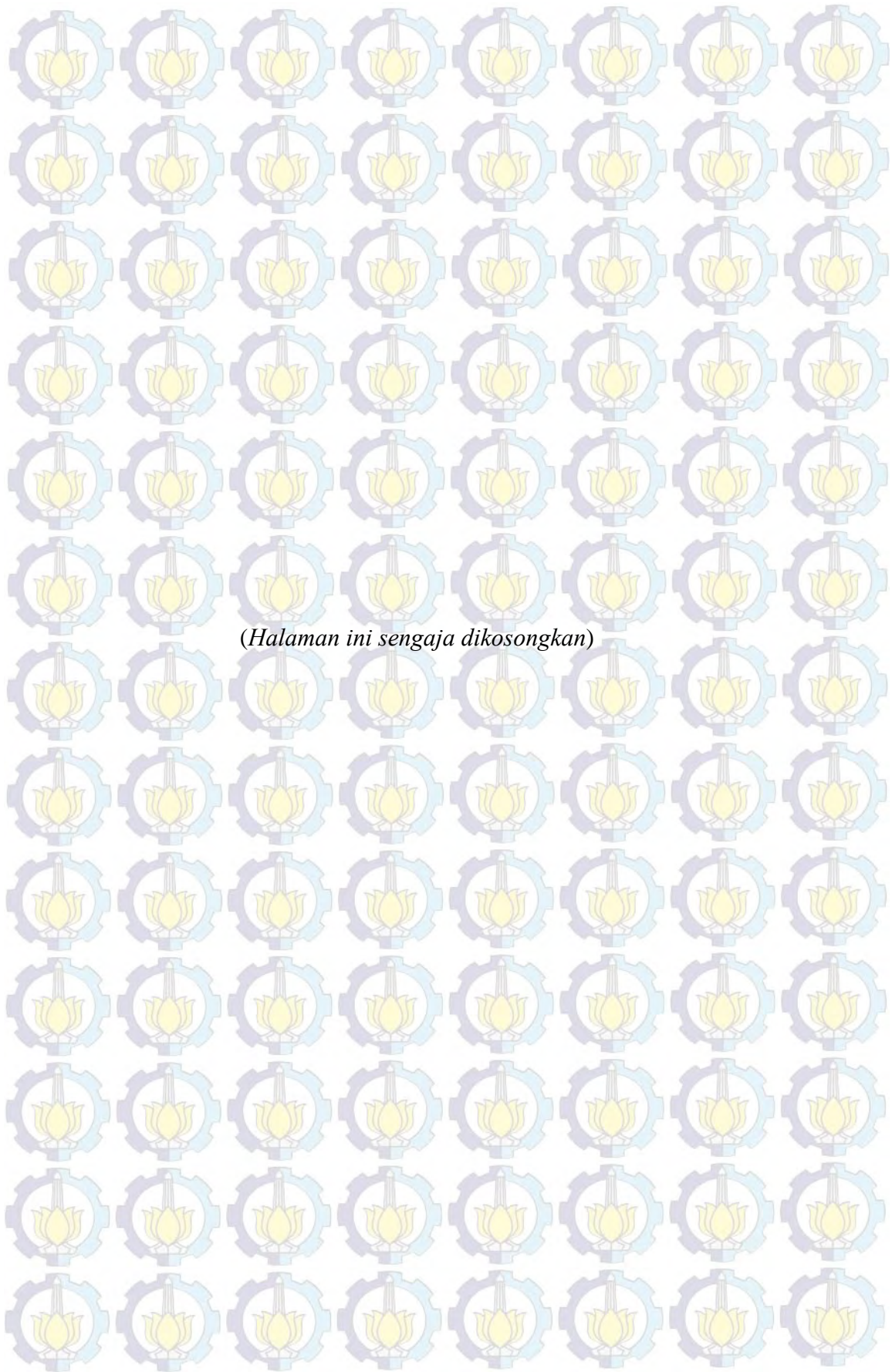
2.3.3. Matriks Penimbang Spasial.....	15
2.4. Model <i>Mixture</i>	16
2.5. Identifikasi Model <i>Mixture</i> dengan Histogram.....	17
2.6. Model Frailty.....	18
2.7. Analisis Bayesian.....	19
2.7.1. Distribusi Prior.....	20
2.7.2. Likelihood Pada Model <i>Mixture</i>	21
2.7.3. Markov Chain Monte Carlo.....	21
2.7.4. <i>Gibbs Sampling</i>	22
2.7.5. <i>Metropolis-Hasting</i>	22
2.7.6. Reversible Jump Markov Chain Monte Carlo.....	23
2.8. Odds Ratio.....	24
2.9. WinBUGS.....	24
2.10. Angka Lama Sekolah dan Rata-Rata Lama Sekolah.....	25
2.11. Variabel-Variabel Yang Mempengaruhi Angka Lama Sekolah.....	27
BAB III METODOLOGI PENELITIAN.....	29
3.1. Sumber Data.....	29
3.2. Metode Pengambilan Data.....	30
3.3 Variabel Penelitian.....	31
3.3.1. Variabel Respon.....	31
3.3.2. Variabel Prediktor.....	33
3.4. Struktur Data.....	34
3.5. Metode dan Tahapan Penelitian.....	34
3.6. Jadwal Kegiatan Penelitian.....	38

BAB IV HASIL DAN PEMBAHASAN	39
4.1. Model Survival dengan Penambahan Konsep Mixture dan Pertimbangan Adanya Efek Spasial Menggunakan Pendekatan Bayesian.....	39
4.1.1. Hubungan Antara Fungsi <i>Survival</i> dengan Fungsi <i>Hazard</i>	39
4.1.2. Penambahan Efek Random Spasial Pada Model Hazard Proporsional	41
4.1.3. Distribusi Prior dan Joint Distribusi Posterior dari Model Mixture Survival Spasial	43
4.1.4. Estimasi Parameter Model Mixture Survival Spasial Dengan Menggunakan RJMCMC dan Gibbs Sampling.....	47
4.2. Model Mixture Survival Spasial Pada Angka Lama Sekolah Anak Umur 16-18 Tahun Berdasarkan Variabel-Variabel Yang Mempengaruhi.....	47
4.2.1. Karakteristik Responden Anak Umur 16-18 Tahun Berdasarkan Variabel-Variabel Prediktornya.....	48
4.2.2. Pengujian Asumsi Model <i>Hazard</i> Proporsional.....	56
4.2.3. Penimbang Spasial	57
4.2.4. Mengidentifikasi Distribusi Data Waktu Survival Angka Lama Sekolah	58
4.2.5. Menentukan Model dan Parameter Model <i>Mixture</i> <i>Survival</i> Spasial	60
BAB V KESIMPULAN DAN SARAN	77
3.1. Sumber Data	77
3.2. Metode Pengambilan Data	79
DAFTAR PUSTAKA.....	81



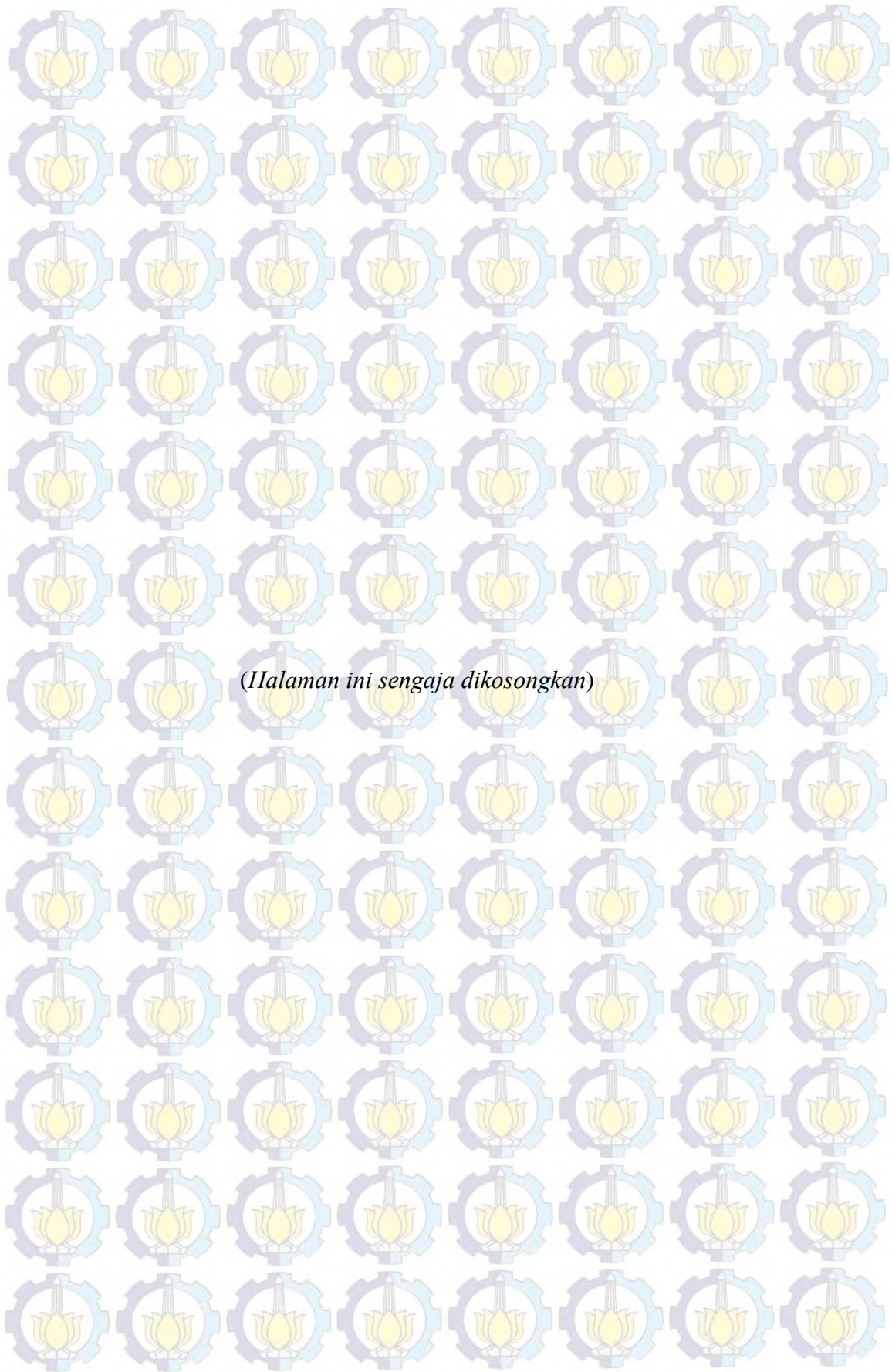
DAFTAR GAMBAR

Gambar 2.1. Ilustrasi Persinggungan (<i>Contiguity</i>)	15
Gambar 3.1. Diagram Alir Penelitian	37
Gambar 4.1. Diagram Batang Variabel Jenis Kelamin Berdasarkan Angka Lama Sekolah Anak Umur 16-18 Tahun.....	50
Gambar 4.2. Pie Chart Jenis Kelamin Responden Anak Umur 16-18 Tahun.....	50
Gambar 4.3. Pie Chart Status Perkawinan dari Responden Anak Umur 16- 18 Tahun	51
Gambar 4.4. Pie Chart Status Bekerja dari Responden Anak Umur 16-18 Tahun	51
Gambar 4.5. Pie Chart Jenis Kelamin Kepala Rumah Tangga	52
Gambar 4.6. Pie Chart Tingkat Pendidikan Terakhir Kepala Rumah Tangga	53
Gambar 4.7. Pie Chart Lokasi Rumah Responden.....	54
Gambar 4.8. Plots $\ln[-\ln S(t)]$ Untuk Uji Proporsional <i>Hazard</i>	55
Gambar 4.9. Histogram Angka Lama Sekolah Anak Umur 16-18 Tahun.....	57
Gambar 4.10. Contoh Pendekatan Distribusi Normal Pada Komponen <i>Mixture</i> Pada Penelitian Green dan Sylvia (1997) Data Enzyme	58
Gambar 4.11. Histogram Data <i>Mixture</i> 1 Sampai <i>Mixture</i> 6.....	59



DAFTAR TABEL

Tabel 3.1. Daftar Variabel Yang Digunakan Dalam Penelitian.....	31
Tabel 3.2. Struktur Data.....	34
Tabel 3.3. Rencana Jadwal Kegiatan Penelitian	41
Tabel 4.1. Deskriptif Responden Anak Berumur 16-18 Tahun	46
Tabel 4.2. Deskriptif Status Sensor Anak Umur 16-18 Tahun Per Kabupaten/Kota Se Jawa Timur.....	47
Tabel 4.3. Deskriptif Rata-Rata Pengeluaran terhadap Status Bekerja Responden	52
Tabel 4.4. Uji Distribusi Waktu <i>Survival</i>	57
Tabel 4.5. Estimasi Parameter Model <i>Mixture Survival</i> Spasial.....	60
Tabel 4.6. Odds Ratio	64
Tabel 4.7. Efek Spasial dari Setiap Kabupaten di Jawa Timur.....	73



BIODATA PENULIS



Penulis dengan nama lengkap Jihadil Qudsi S. ini lahir di Mataram, 18 Februari 1990. Jenjang pendidikan yang telah ditempuh penulis adalah Sekolah Dasar di SDN 7 Mataram pada tahun 1996- 2002, SMP Negeri 2 Mataram pada tahun 2002-2005, dan SMA Negeri 1 Mataram pada tahun 2005-2008. Setelah lulus SMA, penulis mendaftar di Jurusan Matematika Universitas Mataram dengan Jalur PMJK Prestasi dan menyelesaikan studinya pada akhir tahun 2012.

Pada pertengahan tahun 2013, penulis mendaftar di Program Studi Magister Statistika ITS dan tercatat sebagai mahasiswa aktif hingga tahun 2015. Sejak menjadi murid SMA hingga menjadi mahasiswa penulis aktif dalam organisasi-organisasi. Penulis juga gemar olahraga badminton dan sepak bola.

Segala saran dan kritik yang membangun selalu penulis harapkan untuk kebaikan pada masa yang akan datang. Penulis dapat dihubungi melalui jihadil.qudsi@gmail.com.

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Analisis *survival* adalah metode statistika yang mempelajari lamanya suatu proses berjalan hingga kejadian atau peristiwa terjadi yang biasa dikenal dengan nama *failure event*. Waktu ketahanan hidup (*survival time*) atau T merupakan waktu dari awal perlakuan sampai terjadinya respon pertama kali yang ingin diamati. Respon yang dimaksud berupa waktu yang diperlukan hingga suatu peristiwa yang diharapkan terjadi dan mungkin saja belum ditemukan pada saat pengumpulan data berakhir sehingga waktu *survival*-nya tidak dapat diamati. Pada kondisi demikian, pengamatan disebut sebagai pengamatan tersensor (Collet, 1994). Metode regresi *survival* adalah metode regresi yang digunakan untuk melihat variabel-variabel yang menyebabkan terjadinya suatu peristiwa (biasa dikenal dengan nama *time dependent covariate*) dengan variabel responnya adalah waktu ketahanan hidup. Salah satu metode regresi *survival* yang sering digunakan untuk menyelesaikan permasalahan ini adalah regresi *cox proportional hazard*.

Analisis *survival* banyak diterapkan dalam bidang kesehatan atau medis dan dikenal dengan berbagai istilah di bidang lain seperti *event history analysis* dalam bidang sosiologi, analisis *reliability* atau *failure time analysis* dalam bidang teknik dan industri, dan *duration analysis* atau *transition analysis* di bidang ekonomi (Zang, 2008).

Perkembangan analisis *survival* telah banyak dilakukan, antara lain dengan menambahkan konsep distribusi *mixture*. Distribusi *mixture* menggabungkan sejumlah komponen yang kemungkinan berasal dari distribusi yang sama atau bahkan berbeda sehingga dapat memberikan gambaran yang lebih spesifik mengenai sifat-sifat data. Model dengan menggunakan konsep distribusi *mixture* akan memberikan hasil yang lebih akurat karena lebih mendekati distribusi data yang sebenarnya (Marin, Mengersen, dan Robert, 2001). Beberapa penelitian *mixture survival* dilakukan antara lain oleh Ando, Imoto dan Miyano (2004) yang meneliti tentang model *kernel mixture survival* pada beberapa jenis penyakit

kanker; Muthen dan Masyn (2005) yang melakukan penelitian analisis *mixture survival* untuk jenis waktu diskrit (*discret time*) pada kasus kelangsungan sekolah siswa; Zhang (2008) yang melakukan penelitian tentang analisis *mixture survival* melalui pendekatan parametrik dengan menggunakan algoritma EM (*Expectation and Maximization*) yang diaplikasikan pada data simulasi dan data transplantasi hati; dan Hariyanto (2009) yang memodelkan lama pencari kerja di pulau Jawa tahun 2007 dengan metode *mixture survival*.

Telah pula dikembangkan analisis survival dengan penyertaan efek spasial pada analisis survival, dimana efek spasial seringkali mempengaruhi waktu terjadinya suatu kejadian. Efek spasial secara umum dapat dibedakan menjadi dua, yaitu pendekatan *geostatistical* yang menggunakan lokasi geografis dan pendekatan *lattice* yang menggunakan posisi dari suatu daerah relatif terhadap daerah yang lain. Penyertaan efek spasial pada analisis survival, oleh beberapa peneliti menambahkan model efek random untuk mengatasi heterogenitas/sumber-sumber variansi yang tidak dijelaskan dalam model (Darmofal, 2008). Penelitian tersebut antara lain dilakukan dalam penelitian Berry dan Starr (1990 dan 1991), yang pertama kali mengenalkan efek random (*frailty*) yang disertakan dalam model *survival* dimana pembobot untuk kebergantungan spasial dinyatakan dalam jumlah atau proporsi dari daerah-daerah yang saling berdekatan. Selanjutnya Li dan Ryan (2002) meneliti model *spatial survival* dengan pendekatan *semiparametric frailty models* pada data simulasi. Sedangkan Banerjee, Wall dan Carlin (2003), mengembangkan model *survival* hirarkhi yang melibatkan dependensi efek random (*frailty*) pada data spasial, dimana dependensi ini dinyatakan dalam prior *Conditionally Autoregressive* (CAR) yang dikembangkan oleh Besag, York dan Mollie (1991). Prior CAR mengijinkan adanya autokorelasi spasial pada efek random data waktu hingga suatu *event* terjadi pada daerah yang saling berdekatan dan dinyatakan dalam matriks *adjacent*. Salah satu penelitian yang juga menggunakan analisis *survival* spasial dengan efek random (*frailty*) yaitu Hasyim (2012), yang meneliti kasus demam berdarah dengue di kabupaten Pamekasan dengan metode analisis *mixture survival* spasial dengan *frailty* berdistribusi CAR.

Seperti yang telah dituliskan pada bagian awal sub-bab ini, bahwa analisis *survival* sering digunakan pada bidang kesehatan, teknik, industri, sosiologi dan ekonomi. Beberapa penelitian analisis *survival* di atas juga lebih banyak meneliti kasus pada bidang kesehatan. Dalam penelitian ini, peneliti mencoba menggunakan analisis *survival* pada bidang sosiologi yang mencakup bidang pendidikan.

Indikator pendidikan merupakan salah satu indikator dalam membangun Indeks Pembangunan Manusia (IPM), dimana ada dua angka yang digunakan dalam membangun indikator pendidikan yaitu angka persentase warga melek huruf dan angka rata-rata lama sekolah. Di Indonesia perhitungan IPM dimulai sejak tahun 1990 oleh Badan Pusat Statistik (BPS) melalui Survei Sosial Ekonomi Nasional (SUSENAS). Pemerintah telah membuat program sedemikian rupa untuk meningkatkan kedua angka tersebut. Dalam meningkatkan angka melek huruf, pemerintah memberlakukan program belajar membaca dan menulis gratis untuk warga yang sama sekali tidak bisa membaca maupun menulis. Sedangkan guna meningkatkan rata-rata lama sekolah, pemerintah memberlakukan wajib belajar 9 tahun bagi warga setiap warga yang mengenyam pendidikan. Program ini mewajibkan setiap warga negara Indonesia untuk bersekolah selama 9 (sembilan) tahun pada jenjang pendidikan dasar, yaitu dari tingkat kelas 1 Sekolah Dasar (SD) atau yang sederajat hingga kelas 9 Sekolah Menengah Pertama (SMP) atau yang sederajat.

Penelitian ini memfokuskan pada angka rata-rata lama sekolah, dikarenakan sulit untuk meningkatnya angka rata-rata lama sekolah setiap daerah. Angka rata-rata lama sekolah ini disusun atas tingkat pendidikan/angka lama sekolah dari setiap individu atau warga yang sudah berumur 15 tahun ke atas. Jadi dapat dikatakan untuk meningkatkan rata-rata lama sekolah suatu daerah, harus dengan cara meningkatkan angka lama sekolah setiap individu/warga yang tinggal di daerah tersebut.

IPM provinsi Jawa Timur tahun 2012 oleh Badan Pusat Statistik (BPS) Jawa Timur, dari 38 kabupaten/kota di provinsi Jawa Timur hanya 9 kabupaten/kota yang memiliki rata-rata lama sekolah di atas 9 tahun, artinya sebagian besar warga di 29 kabupaten/kota lainnya di provinsi Jawa Timur tidak

lulus atau tidak memiliki ijazah SMP/ sederajat yang formal. Pada tahun yang sama, tahun 2012, tingkat partisipasi sekolah anak umur 16-18 tahun (dalam usia sekolah menengah atas) di Jawa Timur sebesar 61,87 artinya sebesar 61,87% dari total anak yang berumur 16-18 dapat bersekolah (BPS, 2012). Jika melihat pada 3 tahun sebelumnya, tahun 2009, di mana anak yang berumur 16-18 tahun pada tahun 2012, masih berumur 13-15 tahun (dalam usia sekolah menengah pertama), tingkat partisipasi sekolahnya sebesar 88%. Berarti dalam waktu 3 tahun terjadi penurunan tingkat partisipasi sekolah sebesar 26,13% dari jenjang umur 13-15 ke jenjang umur 16-18 tahun.

Fenomena banyaknya anak yang putus sekolah karena himpitan ekonomi sehingga harus berhenti sekolah karena bekerja atau bahkan *drop out* (dikeluarkan dari sekolah), menjadi salah satu penyebab rendahnya tingkat partisipasi sekolah setiap individu sehingga menyebabkan angka lama sekolah rendah juga. Fenomena putus sekolah juga bisa terjadi karena kuota/jumlah sekolah yang kurang ataupun karena kapasitas kursi/daya tampung sekolah yang terbatas yang disediakan oleh pemerintah disetiap jenjang pendidikan yang semakin sedikit. Hal tersebut dapat dilihat dari rasio murid terhadap banyak sekolah menengah pertama dan sederajat yang berstatus negeri dan swasta di Jawa Timur pada tahun ajaran 2012/2013 yaitu 270/1, yang artinya secara umum 1 sekolah hanya bisa menampung maksimal 270 anak setiap tahun ajaran baru. Banyak sekolah negeri dan swasta yang ada di Jawa Timur sebanyak 6.824 sekolah, hal tersebut berarti hanya 1.839.607 anak yang dapat ditampung untuk bersekolah pada tingkat menengah pertama atau yang sederajat setiap ajaran baru, sedangkan banyak anak yang berumur 13-15 tahun pada tahun 2012 sekitar 2.007.866 anak (BPS, 2012).

Penelitian dengan memodelkan angka lama sekolah sendiri sudah pernah diteliti oleh Santoso (2009) dengan menggunakan metode pendekatan *spline multivariable* dan MARS, sedangkan Sulistiyawati (2009) memodelkan angka lama sekolah di kabupaten Boalemo provinsi Gorontalo dengan metode *mixture survival*.

Angka lama sekolah yang membangun rata-rata lama sekolah akan dijadikan variabel respon dan akan dilihat ketahanan (*survive*) lama pendidikan atau angka lama sekolah anak yang berumur 16-18 di provinsi Jawa Timur,

dimana pada usia tersebut masih dalam usia sekolah. Namun penelitian ini akan dihadapkan pada masalah adanya data yang tidak teramati karena terbatasnya waktu penelitian dan kegagalan mengalami kejadian akhir sehingga terdapat data yang tidak lengkap atau data tersensor. Ketika dalam analisis terdapat data tidak lengkap atau tersensor maka dimungkinkan akan menghasilkan interpretasi yang bias (Hobcraft, McDonald, Menken, Rodriguez, dan Trussel 1984). Maka dari itu, penelitian ini menggunakan analisis *survival* dengan tujuan utama untuk menganalisis data yang selalu positif dalam skala pengukuran berupa jarak interval data awal dan akhir yang panjang (Ducrocq, 1997).

Efek spasial dalam penelitian ini karena kebijakan pemerintah daerah pada setiap daerah berbeda-beda dalam menyikapi masalah pendidikan, sehingga perlu pengkajian daerah kabupaten di Provinsi Jawa Timur dimana peran pemerintah, peran masyarakat dan budaya setempat yang disatukan dalam efek spasial, seberapa besar mempengaruhi angka lama sekolah. Maka berdasarkan uraian di atas judul penelitian yang diajukan peneliti yaitu model *mixture survival* spasial pada angka lama sekolah anak umur 16-18 tahun di Jawa Timur tahun 2012.

1.2. Rumusan Masalah

Adapun permasalahan yang diangkat dalam penelitian ini adalah

1. Bagaimana model *survival* dengan penambahan konsep *mixture* dan pertimbangan adanya efek spasial menggunakan pendekatan bayesian?
2. Bagaimana model *mixture survival* spasial pada angka lama sekolah anak umur 16-18 tahun di Jawa Timur berdasarkan variabel-variabel yang mempengaruhi tinggi atau rendahnya angka lama sekolah?

1.3. Tujuan Penelitian

Adapun tujuan penelitian ini antara lain :

1. Memperoleh model *survival* dengan penambahan konsep *mixture* dan pertimbangan adanya efek spasial menggunakan pendekatan bayesian.
2. Memperoleh model *mixture survival* spasial pada angka lama sekolah anak umur 16-18 tahun di Jawa Timur berdasarkan variabel-variabel yang mempengaruhi tinggi atau rendahnya angka lama sekolah.

1.4. Manfaat Penelitian

Adapun manfaat yang bisa diperoleh dari penelitian ini yaitu :

1. Untuk mengembangkan wawasan statistika khususnya bidang analisis *survival*, distribusi *mixture* dan spasial.
2. Hasil penelitian ini diharapkan dapat memberikan masukan kepada pemerintah dalam merancang program dan kebijakan yang terkait dengan bidang pendidikan untuk menanggulangi masalah-masalah yang terkait dengan meningkatkan rata-rata lama sekolah guna untuk meningkatkan nilai IPM di Provinsi Jawa Timur.

1.5. Batasan Permasalahan

Mengacu pada permasalahan di atas, ruang lingkup dalam penelitian ini dibatasi pada beberapa hal, antara lain

1. Sensor yang digunakan dalam penelitian ini adalah sensor kanan (*right censor*) yang berarti bahwa jika responden sampai dengan masa pencacahan selesai belum mengalami *failure event* maka waktunya dibatasi hanya sampai dengan berakhirnya masa pencacahan.
2. Responden yang diambil hanya responden berumur 16-18 tahun yang berstatus sebagai anak pada sebuah rumah tangga.

BAB 2

TINJAUAN PUSTAKA

Pada bab ini untuk tinjauan pustaka dibagi menjadi dua bagian, yaitu tinjauan statistika dan tinjauan non statistika. Tinjauan statistika membahas tentang uji *goodness of fit* analisis *survival*, model *mixture*, model *frailty*, dan analisis bayesian, dan ukuran kelayakan model. Sedangkan tinjauan non statistika membahas tentang angka lama sekolah dan rata-rata lama sekolah serta variabel-variabel yang mempengaruhi angka lama sekolah.

2.1. Uji *Goodness of Fit*

Uji *goodness of fit* distribusi data bertujuan untuk menduga distribusi data. Pada data lama waktu (*lifetime*), uji *goodness of fit* dilakukan pada awal analisis data *survival*. Terdapat beberapa cara dalam melakukan uji *goodness of fit*, diantaranya adalah dengan menggunakan uji Anderson-Darling, Kolmogorov-Smirnov, dan Chi-Square (Law dan Kelton, 2000). Dalam penelitian ini akan digunakan uji *goodness of fit* Anderson-Darling. Hipotesis yang digunakan adalah sebagai berikut:

H_0 : Distribusi data waktu *survival* sesuai dengan distribusi dugaan

H_1 : Distribusi data waktu *survival* tidak sesuai dengan distribusi dugaan

Pengujian hipotesis di atas dilakukan dengan menggunakan statistik uji sebagai berikut,

$$A_n^2 = \left(-\frac{1}{n} \left\{ \sum_{i=1}^n (2i-1) [\ln F(x_i) + \ln(1-F(x_{n+1-i}))] \right\} \right) - n \quad (2.1)$$

dimana F merupakan fungsi distribusi kumulatif dari distribusi yang dihipotesiskan dan x_i merupakan data waktu *survival* yang diurutkan.

Pengambilan keputusan atas hipotesis dengan menggunakan statistik uji (2.1) didasarkan pada kondisi berikut,

Tolak H_0 jika $A_n^2 > a_{n,1-\alpha}$ atau $P_value < \alpha$

dengan nilai $a_{n,1-\alpha}$ adalah nilai tabel dari Anderson-Darling. Semakin kecil nilai Anderson-Darling pada distribusi dugaan, maka data semakin mendekati distribusi dugaan tersebut.

2.2. Analisis Survival

Analisis *survival* adalah salah satu metode statistika untuk menganalisis data dimana variabel responnya berupa waktu sampai suatu *event* terjadi. *Event* didefinisikan sebagai peristiwa ekstrim yang mungkin terjadi pada individu, misalnya kesembuhan maupun kematian setelah dilakukan proses pengobatan. Sedangkan respon *survival* didefinisikan sebagai *range* waktu dari awal penelitian sampai suatu *event* terjadi atau sampai penelitian berakhir, misalnya *range* waktu awal penelitian sampai individu mati atau sampai penelitian berakhir (Kleinbaum dan Klein, 2005). Analisis *survival* banyak diterapkan dalam bidang medis dan biostatistika serta dikenal dengan berbagai istilah di bidang lain seperti: *event history analysis* dalam bidang sosiologi, analisis *reliability* atau *failure time analysis* dalam bidang teknik dan industri, dan *duration analysis* atau *transition analysis* di bidang ekonomi (Cox dan Oakes, 1984). Ada tiga elemen yang harus diperhatikan dalam menentukan waktu *survival* t (Zang, 2008), yaitu :

1. *Time origin/starting point* atau merupakan titik awal, misalnya tanggal lahir pasien atau tanggal dimulainya suatu pengobatan.
2. *Failure time* yaitu waktu berakhirnya *failure event*, misalnya tanggal kematian atau tanggal keluar dari rumah sakit (telah dinyatakan sembuh).
3. *Measurement scale of time* atau skala pengukuran waktu, misalnya skala hari atau bulan.

Dalam menganalisis dengan tipe data lama waktu tidak dapat menggunakan analisis regresi linier biasa karena terdapat data tersensor dan tidak pula dapat menggunakan analisis regresi logistik karena lamanya waktu data yang diikuti tidak sama, sehingga yang paling sesuai adalah menggunakan analisis *survival* (Vittinghof, Glidden, Shiboski, dan McCulloch, 2005). Analisis *survival* memiliki ciri khusus, yaitu distribusi data lama waktu (*lifetime*) bersifat menceng atau *skew* kanan karena nilainya akan selalu positif dan karena adanya data tersensor (Lee,

1992). Data tersensor merupakan kemungkinan beberapa individu tidak bisa diobservasi sampai terjadinya *failure event* (Miller, 1998).

Secara umum, ada tiga alasan terjadinya penyensoran (Collet, 1994), diantaranya adalah sebagai berikut.

1. *Lost of follow up* yaitu jika obyek pengamatan meninggal ataupun pindah.
2. *Drop out* yaitu jika perlakuan harus dihentikan karena suatu alasan tertentu misalnya pemberian kemoterapi yang dihentikan karena efek buruknya lebih besar dibanding manfaatnya.
3. *Termination of study* yaitu jika masa penelitian berakhir sedangkan obyek pengamatan belum sampai pada *failure event*.

Data tersensor dibagi menjadi dua tipe (Collet, 2003), yaitu:

1. Data tersensor kiri, terjadi bila waktu *survival* sebenarnya seorang individu kurang atau sama dengan waktu *survival* individu saat di observasi.
2. Data tersensor kanan, bila waktu *survival* seorang individu tidak lengkap sampai dengan saat periode *follow-up*.
3. Data Sensor interval, bila terjadi pemutusan dalam pengumpulan data dan obyek penelitian mengalami *failure* antar interval waktu tersebut.

Model *survival* digunakan untuk menjelaskan bagaimana resiko (*hazard*) terjadinya suatu *event* tertentu pada suatu waktu dipengaruhi oleh beberapa *covariate* berdasarkan teori yang menunjang peristiwa tersebut. *Hazard rate* merupakan resiko sesaat suatu unit eksperimen pada suatu waktu tertentu yang bertahan, yaitu tidak mengalami peristiwa yang dimaksud hingga waktu berakhir. *Baseline hazard* merupakan resiko terjadinya suatu *event*/peristiwa tanpa mempertimbangkan adanya efek *covariate*, misalnya *time dependency* suatu peristiwa (Darmofal, 2008).

Pada model semiparametrik Cox, tidak terdapat distribusi parametrik khusus untuk *baseline hazard*-nya. Akibatnya, model Cox lebih mengacu hanya pada penggabungan informasi waktu peristiwa yang diamati dibanding dengan menentukan suatu distribusi tertentu untuk interval terjadinya suatu peristiwa. Model Cox mengacu pada semiparametrik karena meskipun tidak ada bentuk distribusi tertentu yang digunakan untuk *baseline hazard* tetapi resiko terjadinya suatu peristiwa tetap dinyatakan sebagai fungsi dari *covariate*. Kelebihan lain

model semiparametrik *Cox* pada fleksibilitas model (berbagai bentuk *baseline hazard*), (Aksioma, 2011). Dikarenakan kelebihan tersebut, dalam penelitian ini yang digunakan model semiparametrik *Cox*.

2.2.1. Fungsi *Survival* dan Fungsi *Hazard*

Misalkan T adalah variabel acak non negatif yang menggambarkan waktu *survival* individu dari suatu populasi. Peluang T pada analisis *survival* secara umum digambarkan ke dalam tiga fungsi yaitu fungsi kepadatan peluang (*Probability Density Function*), fungsi *survival*, dan fungsi *hazard*.

Jika T melambangkan waktu *survival* dan mempunyai distribusi peluang $f(t)$ maka fungsi distribusi kumulatif dinyatakan sebagai berikut:

$$F(t) = P(T < t) = \int_0^t f(u) du . \quad (2.2)$$

Le (1997) menyatakan fungsi *survival* $S(t)$ dapat dinyatakan sebagai peluang seorang individu untuk bertahan lebih lama dari suatu waktu t dan dinyatakan melalui persamaan berikut.

$$S(t) = P(T > t) = 1 - P(T \leq t) = 1 - F(t) . \quad (2.3)$$

Sedangkan fungsi *hazard* dijelaskan sebagai reaksi sesaat atau laju kegagalan (*failure*) sesaat ketika mengalami suatu *event* pada waktu ke- t dan dinyatakan sebagai berikut.

$$h(t) = \lim_{\Delta t \rightarrow 0} \left\{ \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t} \right\} , \quad (2.4)$$

Sedangkan Kleinbaum & Klein (2005) menyatakan bahwa fungsi *hazard* menaksir peluang individu mengalami suatu *event* dalam waktu ke- t atau dalam kasus ini merupakan terjadinya kesembuhan dan dinyatakan melalui persamaan berikut.

$$h(t) dt = P(t \leq T < t + \Delta t \mid T \geq t) . \quad (2.5)$$

Kleinbaum & Klein (2005) juga menyatakan hubungan antara fungsi *survival* dan fungsi *hazard* sebagai berikut:

$$H(t) = -\ln S(t) \quad (2.6)$$

dengan : t : Waktu yang diamati

T : Waktu *survival* seorang individu

$F(t)$: Fungsi distribusi kumulatif dari distribusi data

$f(t)$: Fungsi kepadatan peluang (pdf)

$S(t)$: Fungsi *survival*

$h(t)$: Fungsi *hazard* proporsional

2.2.2. Model dan Asumsi Hazard Proporsional

Nilai variabel prediktor pada model *hazard* proporsional dinyatakan oleh vektor \mathbf{x} , dimana $\mathbf{x} = (x_1, x_2, \dots, x_p)$. Fungsi *baseline hazard* dinyatakan sebagai $h_0(t)$ merupakan fungsi *hazard* untuk tiap-tiap individu dimana semua variabel prediktor dalam vektor \mathbf{x} bernilai 0 (Collet, 2003).

Fungsi *hazard* untuk obyek ke- i dapat ditulis sebagai:

$$h_i(t, \mathbf{x}) = \psi(x_i)h_0(t), \quad (2.7)$$

dengan $\psi(x_i)$ adalah fungsi dari vektor variabel prediktor untuk obyek ke- i .

Fungsi $\psi(x_i)$ dapat diinterpretasikan sebagai fungsi resiko seseorang pada waktu ke- t dengan vektor variabel prediktor x_i relatif terhadap resiko dari suatu obyek yang mempunyai $x = 0$. Secara umum, oleh Le (1997), model *hazard* proporsional diberikan dalam persamaan (2.8) berikut,

$$h(t, \mathbf{x}) = h_0(t) \exp(\beta^T \mathbf{x}), \quad (2.8)$$

dapat dituliskan

$$h(t, \mathbf{x}) = h_0(t) \exp(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p). \quad (2.9)$$

Persamaan (2.8) disebut juga sebagai regresi Cox.

Berdasarkan hubungan fungsi *survival* dengan fungsi *hazard* pada persamaan (2.6), maka menurut Kleinbaum & Klein (2005) dapat dituliskan :

$$S(t, \mathbf{x}) = [S_0(t)]^{\exp(\beta^T \mathbf{x})}, \quad (2.10)$$

$$\ln[S(t, \mathbf{x})] = \exp(\beta^T \mathbf{x}) \ln[S_0(t)]. \quad (2.11)$$

Hal yang menarik pada regresi *cox* adalah data tidak harus memenuhi distribusi apapun (Hosmer dan Lemenshow, 1999). Sedangkan asumsi pemodelan yang harus dipenuhi dalam regresi *cox* yaitu asumsi *hazard* proporsional yang

berarti fungsi *hazard* harus proporsional setiap waktu karena regresi *cox* tidak mengakomodasi variabel yang berubah-ubah sepanjang waktu (Collet, 2003). Proporsional berarti bahwa variabel prediktor independen terhadap waktu dan hubungan antara *hazard* kumulatif sudah proporsional setiap waktu. Asumsi proporsional tersebut dapat diketahui dengan melihat pola plot $-\ln [-\ln S(t)]$ atau $\ln [-\ln S(t)]$ terhadap waktu *survival* untuk setiap variabel prediktor dengan skala kategorik. Asumsi *hazard* proporsional terpenuhi jika pola plot antar kategori dalam variabel prediktor membentuk pola yang sejajar (Kleinbaum dan Klein, 2005). Sedangkan untuk variabel prediktor dengan skala ratio tidak memiliki asumsi apapun.

2.3. Model Survival Spasial

Banerjee *et al.* (2003) mengelompokkan data waktu hingga terjadinya suatu peristiwa (*time-to-event data*) ke dalam strata-strata/kelompok-kelompok, seperti wilayah geografis atau daerah bencana. Dalam kasus ini, pendekatan model hierarki menggunakan *stratum-specific frailties* seringkali cocok. Hal tersebut pertama kali diperkenalkan oleh Vaupel *et al.* (1979) dalam Banerjee *et al.* (2003) dimana terdapat *mixed model* dengan efek random (*frailties*) yang mewakili status kesehatan tiap kelompok.

Misalkan t_{ij} menyatakan waktu hingga terjadinya suatu *event* atau waktu hingga suatu individu i (dimana $i = 1, 2, \dots, n$) dalam strata/kelompok j (dimana $j = 1, 2, \dots, m$) mengalami kejadian lulus, sedangkan \mathbf{x}_{ij} menyatakan vektor dari *covariate* yang mempengaruhi kejadian tersebut, maka asumsi *hazard* proporsional $h(t_{ij}; \mathbf{x}_{ij})$ memungkinkan untuk membentuk model semiparametrik *cox* dengan *hazard rate*,

$$h(t_{ij}; \mathbf{x}_{ij}) = h_0(t_{ij}) \exp(\beta' \mathbf{x}_{ij}), \quad (2.12)$$

dimana *baseline hazards* (h_0) hanya dipengaruhi oleh hasil perkalian dari bentuk eksponensial yang menyertakan *covariate*. Sedangkan untuk persamaan (2.11) menjadi

$$\ln[S(t_{ij}; \mathbf{x}_{ij})] = \exp(\beta' \mathbf{x}_{ij}) \ln[S_0(t_{ij})]. \quad (2.13)$$

Model *Cox* tersebut tidak memuat *intercept* karena tidak dilakukan pengukuran parameter *baseline hazard* (Box dan Jones, 2004 dalam Darmofal, 2008). Dalam model yang menyertakan *frailty*, maka persamaan (2.11) kemudian dapat diperluas menjadi,

$$h(t_{ij}; \mathbf{x}\boldsymbol{\beta}) = h_0(t_{ij}) \exp(\beta^T \mathbf{x}_{ij} + W_j), \quad (2.14)$$

sedangkan persamaan (2.13) menjadi,

$$\ln[S(t_{ij}; \mathbf{x}\boldsymbol{\beta})] = \exp(\beta^T \mathbf{x}_{ij} + W_j) \ln[S_0(t_{ij})], \quad (2.15)$$

apabila dituliskan dalam bentuk seperti persamaan (2.10), maka menjadi

$$S(t_{ij}; \mathbf{x}_{ij}) = [S_0(t_{ij})]^{\exp(\beta^T \mathbf{x}_{ij} + W_j)}, \quad (2.16)$$

dimana W_j merupakan bentuk *stratum-specific frailty* yang dibentuk untuk menyatakan perbedaan antar strata. Secara umum W_j diasumsikan identik, independen dan berdistribusi normal dengan mean bernilai nol dan varian σ^2 , dapat dituliskan sebagai berikut,

$$W_j \stackrel{iid}{\sim} N(0, \sigma^2).$$

Ketika $\sigma^2=0$ maka model (2.14) tereduksi menjadi model (2.12). Pada kenyataannya, σ^2 seperti halnya β dan h_0 , seringkali di estimasi dari data. Akan tetapi menurut Hougaard (2000) dalam Banerjee *et al.*, (2003), distribusi non-normal juga sering digunakan untuk memodelkan *frailties*. Artinya tidak harus distribusi normal untuk asumsi W_j .

Model pada pendekatan spasial *survival* dibentuk dari data *survival* yang tersusun secara spasial, artinya *frailties* W_j dari kelompok/strata yang saling berdekatan menggambarkan kemungkinan bahwa kelompok/strata tersebut memiliki nilai/karakteristik yang serupa, misalnya kelompok kabupaten di sebuah provinsi (Banerjee *et al.*, 2003 dan Darmofal, 2008). Susunan spasial dari kelompok-kelompok tersebut dapat dimodelkan dalam beberapa cara, tetapi secara umum terdapat 2 cara yaitu,

1. Pendekatan geostatistik, yaitu dengan menggunakan lokasi geografis (lintang dan bujur) dari kelompok/strata tersebut.
2. Pendekatan *lattice*, yaitu dengan menggunakan posisi relatif suatu strata/kelompok terhadap kelompok/strata yang lain (*neighboring*).

2.3.1. Model Geostatistik

Untuk memodelkan hubungan spasial antar pengamatan pada lokasi-lokasi tertentu yang telah ditentukan, seringkali digunakan pendekatan klasik sebagaimana dijelaskan oleh Cressie (1993) sebagai model geostatistik. Model tersebut mengasumsikan bahwa proses random yang diamati $Y(s)$ merupakan fungsi kontinu dari s di daerah D (daerah geografis yang diteliti). Model tersebut seringkali digunakan untuk memprediksi nilai suatu pengamatan yang tidak teramati pada beberapa target lokasi.

2.3.2. Model Lattice

Model ini menggunakan informasi jarak dari ketetanggaan (*neighborhood*) atau kedekatan antara satu *region* dengan *region* yang lain. Tobler dalam Anselin (1988) merumuskan hukum *first law of geography* yang berbunyi “segala sesuatu saling berkaitan satu sama lain, namun wilayah yang lebih dekat cenderung akan memberikan efek yang lebih besar daripada wilayah yang lebih jauh jaraknya”.

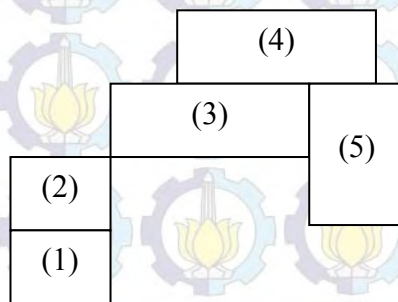
Model ini mengganti distribusi *frailty* yang mengasumsikan proses random \mathbf{W}^* yang dinyatakan melalui indeks kontinu pada daerah D dengan model yang mengasumsikan bahwa \mathbf{W}^* didefinisikan oleh daerah-daerah dengan indeks diskret dimana daerah-daerah tersebut merupakan partisi/bagian dari daerah D . Partisi-partisi tersebut disebut sebagai “*lattice*”. Model ini menggunakan metode penggabungan informasi tentang wilayah-wilayah yang saling berdekatan/ber-tetangga dibandingkan dengan informasi jarak metrik (Banerjee *et al.*, 2003). Akibatnya, distribusi dari efek random \mathbf{W}^* didefinisikan sebagai,

$$\mathbf{W}^* | \lambda \sim \text{CAR}(\lambda)$$

yaitu model *conditionally autoregressive* yang menyatakan adanya ketergantungan spasial pada susunan kovarian, dimana $\lambda = 1/v$ yang merupakan parameter distribusi CAR yang menyatakan presisi atau inverse dari varians distribusi efek randomnya (Wall, 2004).

2.3.3. Matriks Penimbang Spasial

Konsep “*lattice*” merupakan salah satu cara untuk memperoleh matriks penimbang (pembobot) spasial dengan memanfaatkan informasi ketetanggaan dari posisi masing-masing *region* terhadap *region* lainnya. Dalam konteks ketetanggaan (*lattice*) ini, matriks penimbang spasial dari daerah-daerah yang saling bersinggungan dinyatakan melalui indeks diskret (Banerjee *et al.*, 2003). Beberapa metode yang mendefinisikan hubungan kebersinggungan (*contiguity*) antar daerah menurut LeSage (1999) antara lain:



Gambar 2.1 Ilustrasi persinggungan (*Contiguity*)
Sumber : Lesage (1999)

1. *Linear Contiguity* (persinggungan tepi), didefinisikan nilai $w_{ij}=1$ untuk *region* yang berada di tepi (*edge*) kiri maupun kanan *region* yang menjadi perhatian, $w_{ij}=0$ untuk *region* lainnya. Pada gambar 2.1, jika daerah 5 menjadi perhatian, terlihat bahwa $w_{53}=1$, sedangkan yang lain bernilai 0.
2. *Rook Contiguity* (persinggungan sisi), didefinisikan nilai $w_{ij} = 1$ untuk *region* yang bersisian (*common side*) dengan *region* yang menjadi perhatian, dan $w_{ij} = 0$ untuk *region* lainnya. Pada gambar 2.1, jika daerah 3 menjadi perhatian, maka $w_{34}=1$ dan $w_{35}=1$ sedangkan yang lain bernilai 0.
3. *Bishop Contiguity* (persinggungan sudut), didefinisikan nilai $w_{ij} = 1$ untuk *region* yang titik sudutnya (*common vertex*) bertemu dengan sudut *region* yang menjadi perhatian, dan $w_{ij} = 0$ untuk *region* lainnya. Pada gambar 2.1, jika daerah 2 menjadi perhatian, maka $w_{23} = 1$ sedangkan lainnya bernilai 0.
4. *Double Linear Contiguity* (persinggungan dua tepi), didefinisikan nilai $w_{ij} = 1$ untuk dua *entity* yang berada di sisi kiri dan kanan *region* yang menjadi perhatian, dan $w_{ij} = 0$ untuk *region* lainnya.

5. *Double Rook Contiguity* (persinggungan dua sisi), didefinisikan nilai $w_{ij} = 1$ untuk dua *entity* yang berada di sisi kiri, kanan, atas, dan bawah atau dalam peta selatan, utara, barat, dan timur *region* yang menjadi perhatian, dan $w_{ij} = 0$ untuk *region* lainnya.
6. *Queen Contiguity* (persinggungan sisi sudut), didefinisikan nilai $w_{ij} = 1$ untuk dua *entity* yang bersisian (*common side*) atau titik sudutnya (*common vertex*) bertemu dengan *region* yang menjadi perhatian, dan $w_{ij} = 0$ untuk *region* lainnya. Pada gambar 2.1, jika daerah 3 menjadi perhatian, maka $w_{23} = 1$, $w_{34} = 1$ dan $w_{35} = 1$ sedangkan yang lain bernilai 0.

2.4. Model Mixture

Model *mixture* merupakan suatu model yang khas. Model *mixture* akan tampak dari data yang diamati dan data yang ada biasanya terdiri dari beberapa subpopulasi atau grup (McLachlan dan Basford, 1988). Setiap subpopulasi merupakan suatu komponen dari *mixture* tertentu dengan proporsi yang bervariasi untuk setiap komponennya (Carlin, Stern, dan Rubin, 1995). Distribusi *mixture* menggabungkan sejumlah komponen yang kemungkinan berasal dari distribusi yang sama atau bahkan berbeda-beda sehingga dapat memberikan gambaran mengenai sifat-sifat data. Hasil distribusi *mixture* dapat memfasilitasi deskripsi suatu sistem yang kompleks dengan lebih teliti. Distribusi *mixture* menyediakan kerangka parametrik yang fleksibel dalam permodelan dan analisis statistik (Marin *et al*, 2001). Pengembangan awal model *mixture* adalah pada analisis *cluster* yang datanya heterogen namun semakin dikembangkan di berbagai jenis analisis termasuk analisis *survival* (McLachlan dan Peel, 2000). Kasus awal pengembangan model *mixture* adalah berdasarkan penelitian Pearson dari data Weldon dalam McLachlan dan Peel (2000) yang mencoba melakukan *mixture* pada data berdistribusi normal yang berbeda μ_1 dan μ_2 dan variansi σ_1^2 dan σ_2^2 dalam proporsi π_1 dan π_2 . Model *mixture* sendiri telah dikembangkan dengan berbagai pendekatan, yaitu algoritma *Expectation and Maximisation* (EM), *Neural Network*, *Maximum Likelihood*, dan Bayesian.

Misalkan terdapat k komponen dalam sebuah *mixture*, diberikan $f_1(x)$, $f_2(x), \dots, f_k(x)$ sebagai komponen densitas *survival* pertama, kedua sampai K

komponen, maka model *mixture* yang disusun oleh k komponen tersebut akan dapat dituliskan :

$$f(x|\pi) = \pi_1 f_1(x) + \pi_2 f_2(x) + \dots + \pi_K f_K(x) \quad (2.17)$$

dengan π_1 : Nilai proporsi dari komponen *mixture* pertama

π_2 : Nilai proporsi dari komponen *mixture* kedua

π_K : Nilai proporsi dari komponen *mixture* ke- K

sehingga $\pi_1 + \pi_2 + \dots + \pi_K = 1$

Untuk mendeteksi kecenderungan bahwa suatu data berdistribusi *mixture* dapat dilihat dari histogram data tersebut (Iriawan, 2001). Estimasi distribusi *mixture* menggunakan metode *Bayesian* dilakukan dengan menemukan distribusi *posterior*, dengan cara mengalikan distribusi *prior* dan *likelihood* data. Kemudian, estimasi nilai setiap parameter modelnya dapat ditentukan setelah semua *prior* yang relevan telah diberikan (Gamerman, 1997).

Estimasi model *mixture* dengan banyak parameter ini akan memunculkan kesulitan tersendiri. Dimensi model yang besar maupun struktur model yang tinggi akan memberikan kesulitan pada masing-masing langkah estimasi parameternya. Untuk menyelesaikan kesulitan ini, pendekatan *bayesian* memiliki kelebihan dalam penarikan kesimpulan secara numerik, yaitu dengan menggunakan metode *Markov Chain Monte Carlo* (MCMC) (Iriawan, 2000).

2.5. Identifikasi Model *Mixture* dengan Histogram

Metode histogram merupakan metode tertua dan paling sering digunakan dalam estimasi suatu distribusi data. Namun demikian metode ini masih sangat sederhana karena estimasi distribusinya hanya didasarkan pada perkiraan gambar di histogram yang dibentuk. Pembentukan histogram suatu data pengamatan adalah dengan menggunakan titik awal x_0 dan lebar h pada setiap kelas intervalnya, sehingga histogram yang terbentuk berada pada interval $(x_0 + mh, x_0 + (m+1)h)$ dengan m adalah bilangan bulat, sehingga keberadaan kelas-kelas interval sangat dipengaruhi oleh nilai x_0 dan lebar kelas interval h (Silverman, 1986).

Selanjutnya jika dimisalkan dari data pengamatan $\mathbf{x} = (x_1, x_2, \dots, x_n)$, Silverman (1986) mendefinisikan distribusi dari metode histogram adalah sebagai berikut:

$$\hat{f}(x) = \frac{1}{nh} p \quad (2.18)$$

dimana: n = banyaknya data

p = banyaknya data yang ada pada kelas yang sama dengan x

dari persamaan (2.18), tampak bahwa dalam membangun suatu distribusi dengan metode histogram sangat ditentukan oleh banyaknya data, sedangkan banyaknya data dalam setiap interval sangat dipengaruhi oleh keberadaan titik awal x_0 dan lebar kelas h yang digunakan.

Adanya ketergantungan dari titik awal x_0 dan lebar kelas h dalam penentuan distribusi suatu data merupakan kelemahan dari metode histogram. Metode ini dapat digunakan untuk mendeteksi awal bentuk distribusi suatu data secara deskriptif. Selain itu metode histogram juga bisa digunakan untuk menentukan banyak komponen *mixture* disebut cara prediktif, yaitu memperkirakan berapa banyak puncak yang dianggap sebagai banyak komponen *mixture*.

2.6. Model Frailty

Model semiparametrik Cox yang telah dijelaskan sebelumnya mengasumsikan bahwa faktor-faktor yang mempengaruhi *hazard* suatu kejadian telah dijelaskan dalam vektor *covariate* \mathbf{x} . Pada kenyataannya terdapat kasus-kasus tertentu (misalnya kasus dengan data spasial) dimana terdapat keragaman/sumber-sumber variansi yang tidak dapat dijelaskan melalui vektor *covariate* dalam model. Akibatnya, terjadi bias pada pendugaan parameter *survival*. Salah satu cara untuk mengatasi permasalahan tersebut yaitu dengan menyertakan efek random atau *frailty term* dalam model (Darmofal, 2008).

Terdapat 2 pendekatan standar *frailty* dimana penggunaan keduanya bergantung pada kepercayaan peneliti tentang sifat dari keragaman tersebut (Darmofal, 2008).

1. *Unit Specific/Individual frailty terms*, dimana setiap unit dalam pengamatan memiliki *frailty* unik masing-masing.
2. *Hierarchical/Shared frailty terms*, dimana tiap-tiap unit pengamatan saling mengelompok dan unit-unit dalam satu kelompok mempunyai *frailty* bersama sedangkan *frailty* antar kelompok bersifat independen

Pengujian terhadap keragaman yang tidak dapat dijelaskan dalam model secara sederhana melibatkan penaksir parameter variansi dari efek random (θ). Nilai θ positif menyatakan adanya heterogenitas yang tidak dapat dijelaskan dalam model, sebaliknya nilai $\theta = 0$ menyatakan bahwa sumber-sumber variansi telah dapat dijelaskan melalui *covariate* dalam model (Darmofal, 2008). Dengan disertakannya efek random tersebut, dapat ditentukan obyek mana yang dapat bertahan lebih lama (obyek yang cenderung mengalami *failure event*) serta menghindarkan terjadinya bias dan inkonsistensi pada saat melakukan penaksiran parameter sekaligus bias pada standard error yang menyebabkan terjadinya kesalahan inferensi. Dalam konteks spasial *survival* digunakan bentuk *frailty* hirarki atau *stratum-specific frailties* dimana unit pengamatan dikelompokkan dalam wilayah-wilayah *event* yang diteliti.

2.7. Analisis Bayesian

Metode Bayesian ditemukan oleh seorang ilmuwan bernama Thomas Bayes (1702-1761), yang menemukan perlakuan matematika untuk masalah *non trivial* dari inferensi Bayesian, yang digunakan untuk penyelesaian kasus khusus yang kemudian dikenal dengan nama Teorema Bayesian.

Berbeda dengan teori statistika klasik (*frequentist*), analisis bayesian memperlakukan semua parameter yang tidak diketahui sebagai variabel random dan memiliki distribusi (Boldstad, 2007). Teorema Bayesian didasarkan pada distribusi *posterior* yang merupakan perpaduan antara distribusi *prior* (informasi masa lalu sebelum dilakukan observasi) dan data observasi yang digunakan untuk menyusun fungsi *likelihood* (Box dan Tiao, 1973). Hubungan distribusi *posterior* dengan distribusi *prior* dan *likelihood* dapat dituliskan sebagai berikut :

$$\text{Distribusi posterior} \propto \text{likelihood} \times \text{Distribusi prior}$$

Pada teorema Bayes, apabila terdapat parameter θ yang diberikan oleh data observasi y , maka distribusi probabilitas untuk *posterior* θ pada data y akan proporsional dengan perkalian antara distribusi *prior* θ dan fungsi *likelihood* θ yang diberikan oleh data y . Secara matematis dapat ditulis sebagai berikut :

$$f(\theta | x) = \frac{f(x | \theta)f(\theta)}{f(x)} \quad (2.19)$$

atau

$$f(\theta | x) \propto f(x | \theta)f(\theta) \quad (2.20)$$

Dimana $f(\theta|y)$ merupakan distribusi *posterior* yang proporsional dengan perkalian antara fungsi *likelihood* $f(x|\theta)$ dan distribusi *prior* $f(\theta)$.

Spesifikasi dari distribusi *prior* sangat penting dalam metode Bayesian karena *prior* tersebut mempengaruhi bentuk *posterior* yang digunakan sebagai alat pengambilan keputusan. Bila informasi *prior* tersedia, maka informasi untuk distribusi *prior* akan terangkum didalamnya. Namun biasanya informasi *prior* tidak tersedia. Pada kasus ini, perlu penetapan *prior* yang tidak akan mempengaruhi distribusi *posterior* dan “biarkan data berbicara sendiri”. Distribusi tersebut biasa dikenal dengan sebutan *prior* sekawan (*conjugate*) yang parameterisasi distribusi *prior*nya tergolong sebagai *non-informative prior* atau *prior* samar-samar.

2.7.1. Distribusi Prior

Distribusi *prior* merupakan informasi yang terdahulu mengenai parameter (Box dan Tiao, 1973). Berikut beberapa macam distribusi *prior* dalam metode Bayesian :

1. Proper prior atau improper prior (Ntzoufras, 2009)

Penentuan *prior* yang didasarkan pada cara pemberian bobot atau densitas untuk setiap titik disepanjang domain parameter berdistribusi Uniform.

2. Informative prior atau non informative prior (Box dan Tiao, 1973)

Penentuan *prior* yang didasarkan pada ketersediaan pengetahuan atau informasi sebelumnya mengenai pola distribusi data yang diperoleh dari penelitian sebelumnya.

3. Conjugate prior atau non conjugate prior (Box dan Tiao, 1973)

Penentuan *conjugate prior* didasarkan pada pola likelihood dari datanya.

4. *Pseudo prior* (Carlin dan Chip, 1995)

Penentuan *prior* dengan nilai yang disetarakan dengan hasil elaborasi cara *frequentist*.

2.7.2. Likelihood Pada Model Mixture

Fungsi *likelihood* distribusi *mixture* berbeda dengan fungsi *likelihood* distribusi univariat biasa (McLachlan dan Basford, 1988). Jika terdapat data pengamatan sebanyak n yang terdekomposisi ke dalam k kelompok data (subpopulasi) yang masing-masing mempunyai distribusi maka fungsi *likelihood* model *mixture* tersusun dari beberapa *likelihood* data di setiap subpopulasi menurut distribusi masing-masing.

Berdasarkan model *mixture* pada persamaan (2.17), fungsi *likelihood* model *mixture* adalah:

$$l_{mix} = \prod_{i=1}^n f_{mix}(x_i | \theta)$$
$$l_{mix} = \prod_{i_1=1}^{n_1} \pi_1 f(x_{i_1} | \theta_1) + \prod_{i_2=1}^{n_2} \pi_2 f(x_{i_2} | \theta_2) + \dots + \prod_{i_k=1}^{n_k} \pi_k f(x_{i_k} | \theta_k) \quad (2.21)$$

dengan syarat persamaan (2.20) adalah $n_1 + n_2 + \dots + n_k = n$ dan k adalah banyaknya komponen *mixture*.

2.7.3. Markov Chain Monte Carlo (MCMC)

Menurut Carlin (1992), pendekatan MCMC sangat efektif untuk mengurangi beban komputasi dalam menyelesaikan persamaan integrasi yang kompleks dan metode ini memungkinkan proses simulasi dengan mengambil sampel acak dari model stokastik yang sangat rumit.

Ide dasar dari MCMC adalah membangkitkan data sampel dari distribusi *posterior* sesuai proses markov chain dengan menggunakan simulasi monte carlo secara iteratif sehingga diperoleh kondisi yang konvergen terhadap *posterior* (Ntzoufras, 2009). Kondisi tersebut dinamakan kondisi stasioner atau equilibrium. Selanjutnya, sampel parameter dalam markov chain diambil setelah kondisi

stasioner (equilibrium) tercapai. Dengan demikian sampel yang diambil dijamin merupakan sampel dari distribusi *posterior* dari parameter tersebut.

2.7.4. Gibbs Sampling

Gibbs Sampling didefinisikan sebagai suatu teknik simulasi untuk membangkitkan variabel random dari suatu fungsi distribusi tertentu tanpa harus menghitung fungsi densitasnya. Proses Gibbs Sampling dilakukan dengan mengambil sampel dengan cara membangkitkan rangkaian gibbs variabel random berdasarkan sifat-sifat dasar proses Markov Chain. Hal ini merupakan kelebihan dari Gibbs Sampling karena variabel random tersebut dibangkitkan dengan menggunakan konsep distribusi unidimensional yang terstruktur sebagai distribusi full conditional. Gibbs Sampling sangat berguna dalam mengestimasi suatu parameter dalam suatu model yang kompleks yang mempunyai tingkat kerumitan dalam proses integritasi yang kompleks pula dan sulit diselesaikan secara analitis.

2.7.5. Metropolis–Hastings

Metropolis penyusunan *markov chain* dengan menggunakan algoritma *Metropolis – Hastings* dijelaskan oleh Gamerman (1997) sebagai berikut. Dimisalkan terdapat probabilitas transisi Kernel $P(X,Y)$ yang memenuhi persamaan berikut:

$$\pi(X)P(X,Y) = \pi(Y)P(Y,X), \quad \forall (X,Y) \quad (2.22)$$

Selanjutnya probabilitas transisi Kernel $P(X,Y)$ terdiri dari 2 kom ponen, yaitu: $q(x,y)$ dan $\alpha(x,y)$ sehingga $P(X,Y) = q(x,y)\alpha(x,y)$ untuk $X \neq Y$, dimana $q(x,y)$ menotasikan densitas dari X yang akan diajukan untuk mendapatkan *markov chain* berikutnya dan $\alpha(x,y)$ menotasikan probabilitas penerimaan untuk pergerakan dari X ke Y yang dinyatakan sebagai berikut:

$$\alpha(X,Y) = \min \left[\frac{\pi(Y)q(Y,X)}{\pi(X)q(X,Y)} \right] \quad (2.23)$$

adapun $q(x,y)$ menotasikan densitas dari Y untuk kembali ke *state* X .

Jika titik kandidat Y diterima sebagai sampel, maka *state* berikutnya menjadi $X_{t+1} = Y$; sedangkan jika titik kandidat Y ditolak sebagai sampel, maka *markov chain* tidak berubah sehingga $X_{t+1} = X_t$. (Gamerman, 1997) :

Dengan demikian probabilitas transisi Kernel $P(X,Y)$ tidak selalu menghubungkan semua transisi dari X ke Y karena ada kemungkinan penolakan transisi dari X ke Y . Probabilitas terjadinya penolakan transisi dari X ke Y dinyatakan sebagai berikut:

$$P(X,Y) = r(x) = 1 - \int_{-\infty}^{\infty} q(x,y) \alpha(x,y) dy \quad (2.24)$$

2.7.6. *Reversible Jump MCMC*

Pada umumnya penggunaan algoritma MCMC pada proses inferensi bayesian terbatas hanya kasus dimana dimensi vektor parameter model bersifat tetap. Hal ini akan menjadi tidak berlaku pada kondisi dimana dimensi vektor parameter model tidak tetap. Oleh karena itu Green (1995) mengembangkan algoritma Metropolis Hasting untuk mengatasi permasalahan tersebut yang kemudian disebut dengan *Reversible Jump Markov Chain Monte Carlo* (RJMCMC).

Pada pemodelan *mixture* dengan banyak komponen penyusun *mixture* yang tidak diketahui. Richardson dan Green (1997) memanfaatkan algoritma ini untuk mendapatkan model *mixture* yang tepat bagi data yang terindikasi bersifat multimodal. Algoritma ini memanfaatkan proses *birth/death* dan *split/merge* dalam aplikasinya.

Richardson dan Green (1997) menjelaskan proses RJMCMC ini ke dalam enam tipe pergerakan, yaitu :

- | | |
|-----------------------------|--|
| 1. <i>updating</i> ω | 4. <i>updating</i> hyperparameter β |
| 2. <i>updating</i> θ | 5. <i>split/merge</i> komponen-komponen <i>mixture</i> |
| 3. <i>updating</i> z | 6. <i>birth/death</i> dari komponen kosong |

proses sampai dengan langkah ke-4 tidak mengubah dimensi vektor parameter model yang terdiri dari $(\beta, \theta, k, \omega, z)$. Sedangkan langkah 5 dan 6 mengubah banyaknya komponen penyusun *mixture* k satu per satu.

Pergerakan *split/merge* adalah pilihan random antara *split* ($k \rightarrow k+1$) atau *merge* ($k+1 \rightarrow k$). Adapun pergerakan dalam proses *birth/death* merupakan pilihan random antara *birth* ($k \rightarrow k+1$) atau *death* ($k+1 \rightarrow k$).

2.8. Odds Ratio

Odds ratio merupakan suatu ukuran yang digunakan untuk mengetahui tingkat resiko (kecenderungan) yaitu perbandingan antara *Odd* individu dengan kondisi variabel prediktor X pada kategori sukses dengan kategori gagal (Hosmer dan Lemeshow, 1999). Nilai estimasi dari *Odds Ratio* diperoleh dengan mengeksponensialkan koefisien regresi Cox masing-masing variabel prediktor yang signifikan berhubungan dengan *hazard rate*-nya.

Misal X variabel dengan dua kategori yaitu 0 dan 1. Hubungan antara variabel X dan $h(t)$ dinyatakan dengan $h(t|x) = h_0(t)e^{\beta x}$ maka,

- Individu dengan $x = 1$ maka fungsi hazard-nya menjadi $h(t|x = 1) = h_0(t) e^{\beta}$.
- Individu dengan $x = 0$ maka fungsi hazard-nya menjadi $h(t|x = 0) = h_0(t)$.
- Odds Ratio untuk individu dengan $x = 1$ dibanding dengan $x = 0$ yaitu

$$OR = \frac{h(t|x=1)}{h(t|x=0)} = \frac{h_0(t)e^{\beta}}{h_0(t)} = e^{\beta} \quad (2.25)$$

Nilai *OR* artinya bahwa tingkat kecepatan terjadinya *failure event* pada individu dengan kategori $x = 1$ sebesar e^{β} kali tingkat kecepatan terjadinya resiko peristiwa *failure event* pada individu dengan kategori $x = 0$. Sedangkan untuk variabel kontinu, nilai dari e^{β} mempunyai interpretasi sebagai perbandingan *odds ratio* antara individu dengan nilai x lebih besar 1 satuan dibandingkan individu yang lainnya.

2.9. WinBUGS

Winbugs suatu *software* yang dibangun khusus untuk pemodelan yang berdasarkan simulasi stokastik dengan metode Bayesian dengan cara membangkitkan variabel random dari *posterior* distribusi parameter model (Lunn, Thomas, Best dan Spiegelhalter, 2000; Ntzoufras, 2009). BUGS adalah kepanjangan dari *Bayesian inference Using Gibbs Sampling*. Pemodelan dengan menggunakan WinBUGS dapat dilakukan secara *textual* dengan menggunakan

bahasa pemrograman yang sudah dideklarasikan yaitu bahasa BUGS maupun *graphical* dengan menggunakan *DoodleBUGS interface*.

Pengguna yang belum familiar dengan bahasa pemrograman dapat dengan mudah menggambarkan struktur modelnya ke dalam perangkat *doodle* yang merupakan interface dari WinBUGS. Penjelasan mengenai enam subsistem penting untuk program WinBUGS 1.4 adalah sebagai berikut (Lunn dkk., 2000):

1. *Doodle* merupakan tampilan grafis yang digunakan user untuk membangun suatu model regresi dengan menggunakan DAG (*Directed Acyclic Graph*).
2. *Bugs* adalah subsistem yang memiliki beberapa peranan, yaitu mendefinisikan *grammar* untuk spesifikasi model, melakukan validasi spesifikasi model yang telah dibuat, mengumpulkan dan menyimpan representasi model, membentuk pemetaan antara variabel dalam model statistika dengan objek yang mempresentasikan model tersebut, dan mengendalikan algoritma simulasi MCMC.
3. *Samples* adalah subsistem yang menghubungkan detail model yang tersimpan dan monitors yang membentuk *output* berupa teks dan grafis seperti kesimpulan statistik serta gambar densitas *posterior*.
4. *Updater* adalah subsistem yang membentuk objek yang melakukan proses iterasi MCMC berdasarkan data sampel yang diberikan.
5. *Monitors* adalah subsistem yang mendefinisikan tipe objek yang bertanggung jawab untuk menyimpan sampel yang telah terbentuk dari node.
6. *Graph* adalah subsistem yang mengekspor tipe-tipe data dalam suatu node pembangun model grafis.
7. *PKlink* dan *PKBugs* adalah subsistem yang dibangun khusus untuk analisis data *pharmacokinetic*.

2.10. Angka Lama Sekolah dan Rata-Rata Lama Sekolah

Angka lama sekolah yaitu jumlah tahun belajar penduduk usia 15 tahun ke atas yang telah diselesaikan dalam pendidikan formal (tidak termasuk tahun yang mengulang). Untuk menghitung angka lama sekolah dibutuhkan informasi berupa partisipasi sekolah, jenjang dan jenis pendidikan yang pernah/sedang diduduki, ijazah tertinggi yang dimiliki, dan tingkat/kelas tertinggi yang pernah/sedang

diduduki (Sistem Rujukan Statistik BPS). Rumus angka lama sekolah sebagai berikut :

$$ALS = \text{Tahun Konversi} + (\text{Kelas Tertinggi} - 1) . \quad (2.26)$$

Keterangan :

ALS : Angka Lama Sekolah

Tahun Konversi : Tahun Pendidikan yang Ditamatkan (SD = 6 tahun; SMP = 9 tahun; SMA = 12 tahun; D1 = 13 tahun; D2 = 12 tahun; D3 = 13 tahun; D4/S1 = 16 tahun; S2 = 18 tahun; S3 = 21 tahun)

Jika seseorang tidak bersekolah lagi dan tidak memiliki kelas tertinggi maka angka lama sekolahnya adalah tahun konversi tersebut sesuai dengan rumus di bawah ini

$$ALS = \text{Tahun Konversi} \quad (2.27)$$

Sedangkan rata-rata lama sekolah suatu daerah didefinisikan jumlah dari angka lama sekolah dari setiap warga di daerah tersebut dibagi dengan jumlah warga di daerah tersebut dengan rumus di bawah ini. Rata-Rata Lama Sekolah yang menunjukkan rata-rata lamanya bersekolah seseorang dari masuk sekolah dasar sampai dengan tingkat pendidikan terakhir. Pada prinsipnya angka ini merupakan transformasi dari bentuk kategorik tingkat pendidikan tertinggi (TPT) menjadi bentuk numerik. Rata-rata lama sekolah seringkali dijadikan rujukan oleh pemerintah untuk melihat perkembangan pendidikan di daerah tersebut, selain dengan melihat angka melek huruf (Edi, 2012).

$$RLS = \frac{\sum_{i=1}^n ALS_i}{n} . \quad (2.28)$$

Keterangan :

RLS : Rata-rata Lama Sekolah suatu daerah

ALS : Angka Lama Sekolah

n : Jumlah penduduk suatu daerah

Jika rata-rata lama sekolah mengindikasikan untuk melihat kualitas penduduk suatu daerah secara keseluruhan dalam mengenyam pendidikan formal.

Angka lama sekolah mengindikasikan untuk melihat kualitas individual setiap penduduk dalam mengenyam pendidikan formal.

Lamanya bersekolah merupakan ukuran akumulasi investasi pendidikan individu. Setiap tahun tambahan sekolah diharapkan akan membantu meningkatkan pendapatan individu tersebut. Rata-rata lama bersekolah dapat dijadikan ukuran akumulasi modal manusia suatu daerah. Ukuran ini mengatasi masalah kekurangan estimasi dari TPT yang tidak mengakomodir kelas tertinggi yang pernah dicapai individu. Tetapi, jumlah tahun bersekolah ini tidak mengindahkan kasus-kasus tidak naik kelas, putus sekolah yang kemudian melanjutkan kembali, dan masuk sekolah dasar di usia yang terlalu muda atau sebaliknya. Sehingga nilai dari jumlah tahun bersekolah menjadi terlalu tinggi kelebihan estimasi atau bahkan terlalu rendah (*underestimate*). Satuan data yang digunakan adalah tahun.

Tingginya angka dan rata-rata lama sekolah menunjukkan jenjang pendidikan yang pernah/sedang diduduki oleh seseorang. atau semakin lama/tinggi jenjang pendidikan yang ditamatkannya. Dikarenakan rata-rata lama sekolah berasal dari perhitungan angka lama sekolah, maka untuk meningkatkan rata-rata lama sekolah suatu daerah pemerintah harus menaikkan angka lama sekolah setiap individu di daerah tersebut.

2.11. Variabel-Variabel Yang Mempengaruhi Angka Lama Sekolah

Melihat dan mempelajari dari penelitian sebelumnya tentang angka lama sekolah maupun rata-rata lama sekolah seperti penelitian Schidkrout (1981) dalam Hillary (1995) menyatakan bahwa pendidikan orangtua, pendapatan, dan kekayaan orangtua merupakan faktor sosial ekonomi yang mempengaruhi status anak bersekolah atau tidak bersekolah. Artinya pendidikan dan pendapatan orangtua menentukan apakah seorang anak bersekolah atau tidak, walaupun untuk sekolah gratis, namun biaya untuk baju sekolah dan buku penunjang menjadi salah satu alasan para orangtua untuk tidak menyekolahkan anaknya.

Suendra (1999) dalam penelitiannya menyatakan bahwa faktor yang menghambat seseorang anak untuk menyelesaikan pendidikan dasar 9 tahun antara lain faktor ekonomi, faktor sosial budaya, dan faktor geografis.

Ikana (2005) dalam penelitiannya juga menyatakan bahwa ketika dukungan ekonomi tidak memadai bagi kelangsungan pendidikan anak, maka anak

perempuan sering mengalah pada anak laki-laki dan terpaksa putus sekolah, yang berarti jenis kelamin dari seorang anak bisa menjadi faktor tinggi rendahnya angka lama sekolah seorang anak.

Menurut Dalyono (2008), Lingkungan sosial budaya masyarakat adalah semua orang/manusia yang dapat berpengaruh terhadap kehidupan anak. Pengaruh sosial tersebut dapat dilihat secara langsung maupun tidak langsung. Pengaruh secara langsung, seperti terjadi di dalam pergaulan anak sehari-hari dengan teman sebayanya atau orang lain. Yang tidak langsung dapat terjadi melalui jalur informasi, seperti radio atau televisi. Anak-anak yang dibesarkan di kota pola pikirnya berbeda dengan anak di desa. Pada umumnya anak yang tinggal di kota lebih bersikap aktif dan dinamis, bila dibandingkan dengan anak desa yang selalu bersikap statis dan lamban. Itulah sebabnya, perkembangan dan kemajuan anak yang tinggal di kota jauh lebih pesat daripada anak yang tinggal di desa.

Santoso (2009) dalam penelitiannya mengungkapkan variabel prediktor yang mempengaruhi lama sekolah seseorang yaitu topografi wilayah, umur anak, status pekerjaan anak, pendidikan kepala keluarga, pengeluaran perkapita, banyaknya anggota rumah tangga, dan status desa/kota.

Sedangkan Sulistiyawati (2009) dalam penelitiannya tentang faktor yang mempengaruhi angka lama sekolah di daerah kota dan desa di kabupaten Boalemo, provinsi Gorontalo, menyatakan bahwa, umur, status pekerjaan, pendidikan kepala rumah tangga, dan pengeluaran rumah tangga, adalah faktor-faktor yang mempengaruhi angka lama sekolah untuk responden yang bertempat tinggal di daerah perkotaan, sedangkan faktor jenis kelamin, umur, status pekerjaan, pendidikan kepala rumah tangga, pengeluaran rumah tangga, dan banyaknya anggota rumah tangga.

Berdasarkan penjelasan di atas pada penelitian ini variabel-variabel prediktor yang diduga mempengaruhi angka lama sekolah anak umur 16-18 tahun yaitu jenis kelamin, status perkawinan, status bekerja, tingkat pendidikan kepala rumah tangga, jenis kelamin kepala rumah tangga, jumlah anggota rumah tangga, rata-rata pengeluaran rumah tangga per kapita, dan lokasi tempat tinggal (kota atau desa).

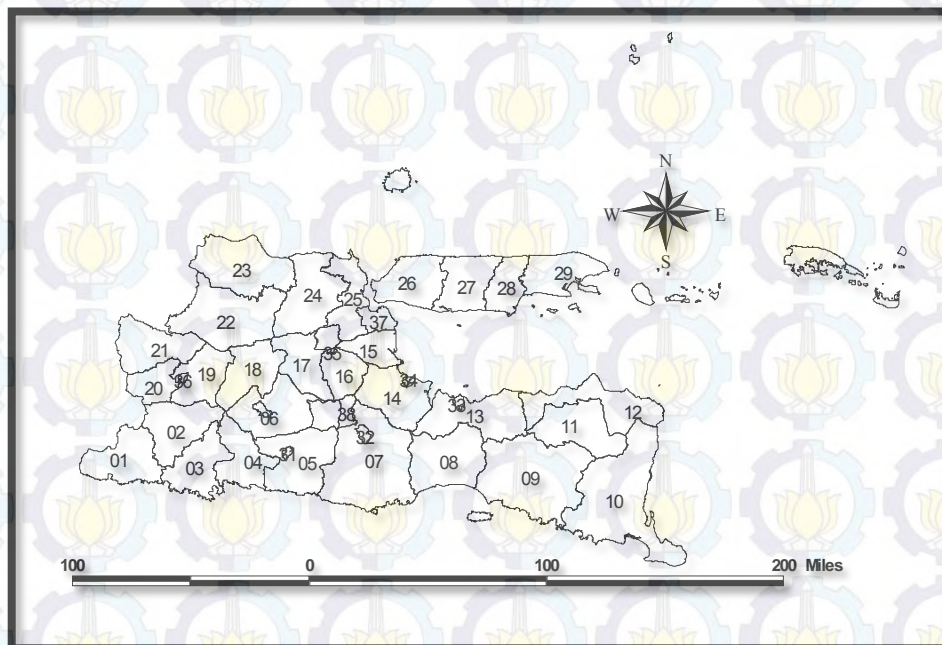
BAB 3

METODOLOGI PENELITIAN

Pada bab ini akan dijabarkan tentang hal mendasar dalam penelitian beserta tahapan-tahapan dalam menjalankannya, diantaranya sumber data, variabel penelitian, metode dan tahapan penelitian serta jadwal penelitian.

3.1. Sumber Data

Data yang dipergunakan dalam penelitian ini data sekunder bersumber dari Badan Pusat Statistik (BPS), yaitu data SUSENAS tahun 2012. Data yang diambil adalah data angka lama sekolah, dan responden yang dinyatakan sebagai *failure event* yaitu apabila lulus jenjang pendidikan sekolah menengah pertama (SMP) atau yang sederajat (MTs, SMPLB dan paket B) baik negeri maupun swasta dan mendapatkan ijazah pada rentang umur 16-18 tahun. Efek spasial dinyatakan dengan kedekatan antara lokasi satu dengan lokasi yang lain (matriks *adjacent*).



Gambar 3.1. Peta Provinsi Jawa Timur

Dimana terdapat 38 kabupaten/kota yang ada pada provinsi Jawa Timur yang bersesuaian dengan kode pada gambar 3.1. yaitu 01 = Pacitan, 02 = Ponorogo, 03 = Trenggalek, 04 = Tulungagung, 05 = Blitar, 06 = Kediri, 07 = Malang, 08 = Lumajang, 09 = Jember, 10 = Banyuwangi, 11 = Bondowoso, 12 = Situbondo, 13 = Probolinggo, 14 = Pasuruan, 15 = Sidoarjo, 16 = Mojokerto, 17 = Jombang, 18 = Nganjuk, 19 = Madiun, 20 = Magetan, 21 = Ngawi, 22 = Bojonegoro, 23 = Tuban, 24 = Lamongan, 25 = Gresik, 26 = Bangkalan, 27 = Sampang, 28 = Pamekasan, 29 = Sumenep, 30 = Kota Kediri, 31 = Kota Blitar, 32 = Kota Malang, 33 = Kota Probolinggo, 34 = Kota Pasuruan, 35 = Kota Mojokerto, 36 = Kota Madiun, 37 = Kota Surabaya, 38 = Kota Batu,

3.2. Metode Pengumpulan Data

Susenas merupakan survei dengan unit sampel rumah tangga yang dilaksanakan oleh BPS dengan lingkup nasional. Susenas 2012 menggunakan tiga jenis kerangka sampel dengan urutan tahapan, yaitu: kerangka sampel untuk pemilihan kecamatan, kerangka sampel untuk pemilihan blok sensus dari kecamatan terpilih, dan kerangka sampel untuk pemilihan sampel rumah tangga dari blok sensus terpilih.

Setiap rumah tangga yang terpilih dalam survei ini dikunjungi oleh petugas pencacah yang bertugas mewawancarai responden sesuai dengan daftar pertanyaan dalam kuisisioner yang sudah disiapkan. Wawancara dilakukan langsung terhadap kepala rumah tangga atau kepada anggota rumah tangga yang dianggap mengetahui keadaan di rumah tangga tersebut. Petugas yang melakukan pengumpulan data adalah Koordinator Statistik Kecamatan (KSK) atau staf BPS kabupaten/kota yang sudah dilatih terlebih dahulu dengan konsep dan definisi serta simulasi survei. Setiap petugas survei juga dibekali dengan instrumen untuk keperluan pengumpulan data, seperti kuesioner dan buku pedoman.

Adapun jumlah sampel terpilih dalam Susenas tahun 2012 di Provinsi Jawa Timur adalah sebanyak 29.274 sampel rumah tangga di Provinsi Jawa Timur, namun yang digunakan adalah rumah tangga yang memiliki anak yang berumur 16-18 tahun yang berjumlah 3.794 rumah tangga.

3.3. Variabel Penelitian

Variabel respon dan variable-variable prediktor yang diperkirakan mempengaruhi angka lama sekolah yang telah disebutkan dalam bab 2 tersaji dalam tabel berikut :

Tabel 3.1. Daftar Variabel Yang Digunakan Dalam Penelitian

Variabel	Uraian	Skala Data	Kategori
(1)	(2)	(3)	(4)
Respon			
Y	Angka Lama Sekolah (tahun)	Rasio	
Prediktor			
X ₁	Jenis Kelamin	Nominal	1 = Laki-Laki 2 = Perempuan
X ₂	Status Perkawinan	Nominal	1 = Belum Kawin 2 = Kawin
X ₃	Status Bekerja	Nominal	1 = Tidak Bekerja 2 = Bekerja
X ₄	Tingkat Pendidikan Kepala Rumah Tangga (tahun)	Ordinal	1 = Tidak Tamat SD 2 = Tamat SD 3 = Tamat SMP 4 = Tamat SMA 5 = Tamat PT
X ₅	Jenis Kelamin Kepala Rumah Tangga	Nominal	1 = Laki-Laki 2 = Perempuan
X ₆	Rata-Rata Pengeluaran Rumah Tangga per Kapita (Rp per Bulan)	Rasio	
X ₇	Jumlah Anggota Rumah Tangga	Ratio	
X ₈	Lokasi Tempat Tinggal	Nominal	1 = Kota 2 = Desa

Untuk menjaga kesamaan persepsi dalam penelitian ini, diberikan definisi untuk masing-masing variabel sebagai berikut:

3.3.1. Variabel Respon (Y)

Variabel respon yang digunakan dalam penelitian ini adalah data angka lama sekolah di Provinsi Jawa Timur tahun 2012 dari responden yang disimbolkan dengan Y (yang tidak pernah menempuh sekolah, yang sedang bersekolah, dan yang pernah menempuh sekolah tapi kini tidak bersekolah), yang

berumur 16-18 tahun (dalam usia sekolah menengah atas) berada dalam batas periode penelitian, dalam satuan tahun, dengan ketentuan sebagai berikut:

- a. Jika seorang responden pernah bersekolah ataupun masih bersekolah hingga dinyatakan lulus karena mendapatkan ijazah sekolah menengah pertama (SMP) atau yang sederajat (MTs, SMPLB dan paket B) baik negeri maupun swasta dan dalam batas periode penelitian, maka waktu *survival* dikategorikan sebagai data *survival* tidak tersensor.
- b. Jika seorang responden sampai dengan batas periode penelitian, terhitung dari responden tersebut mulai bersekolah (tidak dibatasi kapanpun responden tersebut bersekolah) hingga responden tersebut mengalami hal-hal berikut:
 1. Masih bersekolah pada jenjang SD/sederajat atau SMP/sederajat atau melebihi batas akhir penelitian (tahun 2012),
 2. Jika seorang responden tidak pernah sama sekali bersekolah,
 3. Jika seorang responden keluar dari masa belajar (mengundurkan diri) atau mengalami drop out (dikeluarkan dari sekolah),

maka data *survival* dikatakan data tersensor. Variabel respon kemudian dikategorikan menjadi:

$t=0$, data tersensor jika responden tidak mengalami *failure event*. Misal adanya responden tidak pernah bersekolah sama sekali, keluar dari masa belajar (mengundurkan diri) atau mengalami drop out (dikeluarkan dari sekolah), atau melebihi batas akhir penelitian dan responden masih belum mendapatkan ijazah SMP/sederajat.

$t=1$, data tidak tersensor jika responden mendapatkan ijazah SMP/sederajat yang berada dalam batas periode penelitian.

3.3.2 Variabel Prediktor

Variabel prediktor adalah variabel yang digunakan untuk memprediksi variabel respon. Adapun variabel prediktor yang digunakan pada penelitian ini adalah:

- a. Variabel Jenis Kelamin Responden (X_1), adalah jenis kelamin dari responden. Variabel jenis kelamin dikategorikan 1 untuk perempuan dan 2 untuk laki-laki.
- b. Variabel Status Perkawinan Responden (X_2), yaitu status perkawinan responden saat pencacahan. Status perkawinan yang biasa dikenal adalah

belum kawin, kawin, cerai hidup, atau cerai mati. Responden yang berstatus kawin, cerai hidup, atau cerai mati. Selanjutnya dalam penelitian ini diberikan pengkodean ulang (*recoding*) menjadi kawin.

- c. Variabel Status Bekerja Responden (X_3), yaitu status bekerja responden saat pencacahan. Kategori dari status bekerja adalah tidak bekerja dan bekerja.
- d. Variabel Tingkat Pendidikan Kepala Rumah Tangga (X_4), yaitu tingkat pendidikan tertinggi yang dimiliki oleh seorang kepala rumah tangga. Variabel berskala ordinal, dan pengkategorian variabel ini sebagai berikut : 1 untuk tidak tamat SD, 2 untuk tamat SD, 3 untuk tamat SMP, 4 untuk tamat SMA, dan 5 untuk tamat Perguruan Tinggi (PT).
- e. Variabel Jenis Kelamin Kepala Rumah Tangga (X_5), adalah jenis kelamin dari kepala rumah tangga yang dikategorikan 1 untuk perempuan dan 2 untuk laki-laki.
- f. Rata-Rata Pengeluaran Rumah Tangga per Kapita per Bulan (X_7), yaitu rata-rata pengeluaran untuk keperluan sehari-hari yang dikeluarkan oleh sebuah rumah tangga dalam satuan per kapita setiap bulan.
- g. Variabel Jumlah Anggota Rumah Tangga (X_6), adalah jumlah anggota rumah tangga yang berada dalam 1 rumah.
- h. Variabel Lokasi Tempat Tinggal Responden (X_8), lokasi tempat tinggal responden, yaitu di perkotaan atau di pedesaan.

3.4. Struktur Data

Struktur data dari penelitian ini dengan variabel-variabel yang bisa dilihat pada poin sebelumnya disajikan pada Tabel 3.2 berikut :

Tabel 3.2. Struktur Data

No	Y (Angka Lama Sekolah)	T (Status Sensor)	X_1 JK	X_2 SP	X_3 SB	X_4 TPK RT	X_5 JKK RT	X_6 Peng	X_7 JA RT	X_8 LT T	Kab /Kota
1	y_1	t_1	x_{11}	x_{21}	x_{31}	x_{41}	x_{51}	x_{61}	x_{71}	x_{81}	1
2	y_2	t_2	x_{12}	x_{22}	x_{32}	x_{42}	x_{52}	x_{62}	x_{72}	x_{82}	1
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
n	y_n	t_n	x_{1n}	x_{2n}	x_{3n}	x_{4n}	x_{5n}	x_{6n}	x_{7n}	x_{8n}	m

3.5. Metode dan Tahapan Penelitian

Berdasarkan tujuan penelitian yang telah dijelaskan pada bab sebelumnya, maka langkah-langkah analisis dijelaskan sebagai berikut:

1. Mengkaji model *survival* dengan penambahan konsep *mixture* dan pertimbangan adanya efek spasial menggunakan pendekatan bayesian. Langkah-langkahnya antara lain:
 - 1.1. Mengkaji hubungan antara fungsi *survival* dalam persamaan (2.3) dengan fungsi *hazard* dalam persamaan (2.4) dan (2.5) sehingga diperoleh fungsi *hazard* proporsional.
 - 1.2. Menambahkan efek random spasial (*spatial frailty term*) pada model *hazard* proporsional.
 - 1.3. Menentukan distribusi *prior* dan *joint* distribusi *posterior* dari model *mixture survival* spasial berdasarkan distribusi data.
 - 1.4. Menentukan jumlah komponen *mixture* dan estimasi parameter model *mixture survival* spasial dengan menggunakan RJMCMC dan Gibbs Sampler.
2. Menentukan model *mixture survival* spasial pada angka lama sekolah anak umur 16-18 tahun di Jawa Timur berdasarkan faktor-faktor yang mempengaruhi tinggi atau rendahnya angka lama sekolah anak umur 16-18 tahun. Langkah-langkahnya dijelaskan sebagai berikut:
 - 2.1 Mengkaji karakteristik responden anak umur 16-18 tahun berdasarkan jenis kelamin, umur, angka lama sekolah kepala rumah tangga, jenis kelamin kepala rumah tangga, status perkawinan, status bekerja, dan rata-rata pengeluaran rumah tangga per kapita, yaitu dengan melakukan analisis statistik deskriptif meliputi tabel deskriptif, *pie chart*, dan diagram batang.
 - 2.2 Melakukan pengujian asumsi pemodelan *hazard* proporsional menggunakan plot $-\ln[-\ln S(t)]$ atau $\ln[-\ln S(t)]$ terhadap waktu *survival t*.

2.3 Menentukan penimbang spasial

Untuk menentukan penimbang spasial dilakukan dengan cara input peta wilayah provinsi Jawa Timur ke dalam program GeoDa. Kemudian mencari matriks kedekatan (matriks *adjacent*) dengan memasukkan peta *jatim.shp* yang sudah ada pada program GeoDa.

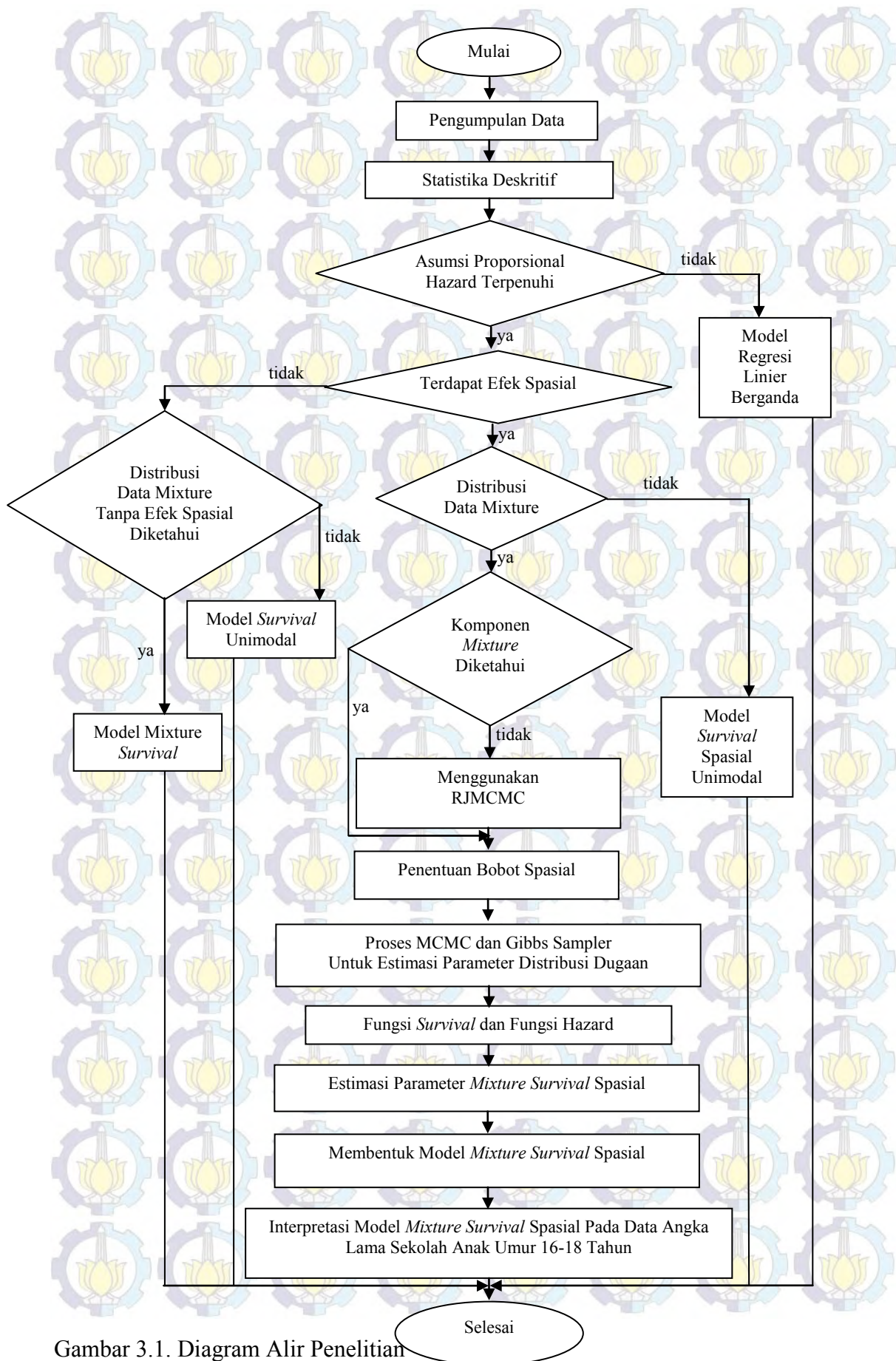
2.4 Mengidentifikasi dan menguji distribusi data waktu *survival* angka lama sekolah dengan menggunakan histogram.

2.5 Jika secara histogram ada indikasi data berdistribusi *mixture* dan banyak komponen tidak diketahui, maka digunakan RJMCMC untuk penentuan komponennya.

2.6 Menentukan model dan parameter model spasial *survival* menggunakan simulasi MCMC dengan tahapan sebagai berikut:

1. Menentukan parameter dari distribusi *mixture* untuk penentuan fungsi *survival* dan fungsi hazard.
2. Menentukan model dan parameter model *survival*.
3. Menentukan matriks penimbang spasial (*weighted matrix*) \mathbf{W}^s .
4. Menentukan distribusi efek random spasial (*spatial frailty term*).
5. Menentukan mean dan varians dari distribusi efek random spasial.
6. Mengestimasi parameter model *survival* dengan efek random spasial σ , β , dan λ serta proporsi dari setiap komponen *mixture* (π) untuk komponen *mixture* melalui simulasi MCMC dengan *Gibbs sampling*.
7. Membentuk model *mixture survival* spasial berdasarkan parameter model yang diperoleh dan akan terbentuk k model *hazard proporsional* yaitu model untuk komponen *mixture* pertama hingga model untuk komponen *mixture* ke- k . Kemudian menggabungkan model-model tersebut ke dalam sebuah model *mixture* yang terdiri dari model komponen *mixture* dengan proporsi (π) masing-masing.
8. Interpretasi model *mixture survival* spasial untuk masing-masing faktor-faktor yang berpengaruh signifikan terhadap laju kelulusan responden (*hazard rate*) pada setiap komponen *mixture* dan efek spasial dari setiap daerah kabupaten/kota.

Untuk lebih jelasnya dapat dilihat dalam diagram alir Gambar 3.1,



Gambar 3.1. Diagram Alir Penelitian

BAB 4

HASIL DAN PEMBAHASAN

Pada bab ini akan dijabarkan tentang hal mendasar dalam penelitian beserta tahapan-tahapan dalam menjalankannya, diantaranya sumber data, variabel penelitian, metode dan tahapan penelitian serta jadwal penelitian.

4.1. Model *Survival* Dengan Penambahan Konsep *Mixture* dan Pertimbangan Adanya Efek Spasial Menggunakan Pendekatan Bayesian.

Pada bagian ini akan menjawab tujuan pertama penelitian yaitu mengenai penjabaran model *survival* dengan penambahan konsep *mixture* dan pertimbangan adanya efek spasial menggunakan pendekatan bayesian, yang meliputi hubungan model *survival* dan model *hazard* serta didapatnya mode *hazard* proporsional, penambahan efek random (*frailty*) dalam model *hazard* proporsional, penambahan efek random spasial (*spatial frailty*) dalam model *hazard* proporsional, penentuan distribusi *prior* dan *joint* distribusi *posterior* dari model *mixture survival* spasial berdasarkan distribusi data, serta estimasi parameter model *mixture survival* spasial menggunakan MCMC dan *gibbs sampling*.

4.1.1. Hubungan Antara Fungsi *Survival* Dengan Fungsi *Hazard*

Dalam analisis *survival*, terdapat dua fungsi yang digunakan yaitu fungsi *survival* dan fungsi *hazard*. Pada bab 2 telah dituliskan bahwa Le (1997) mendefinisikan fungsi *survival* $S(t)$ sebagai probabilitas seorang individu bertahan lebih lama dari waktu t . Fungsi *survival* dapat dilihat pada persamaan (2.3).

Lee (1992) mendefinisikan fungsi *hazard* sebagai peluang kegagalan individu untuk bertahan selama interval waktu yang sangat pendek dengan asumsi bahwa individu tersebut telah bertahan pada awal interval atau limit peluang individu gagal bertahan dalam sebuah interval waktu yang sangat pendek, yaitu dari t sampai Δt jika individu tersebut telah bertahan sampai waktu t seperti yang dituliskan pada persamaan (2.4). Semakin besar nilai *hazard* mengindikasikan

bahwa risiko kegagalan yang dialami individu dalam penelitian semakin tinggi sehingga kemampuan bertahanannya semakin kecil.

Hubungan antara fungsi *survival* dan fungsi *hazard* dapat diperoleh dengan menggunakan teori probabilitas bersyarat, yaitu $P(A|B) = \frac{P(AB)}{P(B)}$, dengan pembilang dalam limit persamaan (2.4) dapat ditentukan hubungannya sebagai berikut:

$$\frac{P(t \leq T < t + \Delta t)}{P(T \geq t)} = \frac{F(t + \Delta t) - F(t)}{S(t)} \quad (4.1)$$

dengan $F(t)$ merupakan fungsi distribusi dari T maka persamaan (2.4) dapat dituliskan:

$$h(t) = \lim_{\Delta t \rightarrow 0} \left\{ \frac{F(t + \Delta t) - F(t)}{\Delta t} \right\} \frac{1}{S(t)} \quad (4.2)$$

dengan mengambil $F'(t) = \lim_{\Delta t \rightarrow 0} \left\{ \frac{F(t + \Delta t) - F(t)}{\Delta t} \right\}$ yang merupakan derivatif pertama dari $F(t)$ maka diperoleh hubungan antara fungsi *survival* dan fungsi *hazard* sebagai berikut:

$$h(t) = \frac{f(t)}{S(t)} \quad (4.3)$$

Diketahui bahwa $F(t) = 1 - S(t)$ dapat dituliskan sebagai $\int f(t)dt = 1 - S(t)$. Jika keduanya diturunkan terhadap t maka diperoleh $f(t) = \frac{d(1 - S(t))}{dt}$ dan nilai $h(t)$ menjadi:

$$h(t) = \frac{\left(\frac{d(1 - S(t))}{dt} \right)}{S(t)} = \frac{\left(-\frac{d}{dt} S(t) \right)}{S(t)} \quad (4.4)$$

$$-h(t)dt = \frac{d(S(t))}{S(t)}$$

dengan mengintegrasikan persamaan (4.4), diperoleh

$$-\int_0^t h(t)dt = \int_0^t \frac{1}{S(t)} d(S(t))$$

$$-\int_0^t h(t)dt = \ln S(t) \Big|_0^t = \ln S(t) - \ln S(0) = \ln S(t)$$

$$S(t) = \exp \left[-\int_0^t h(t)dt \right] \quad (4.5)$$

dengan fungsi kumulatif *hazard* adalah sebagai berikut:

$$H(t) = \int_0^t h(t)dt \quad (4.6)$$

sehingga didapatkan hubungan antara fungsi kumulatif *hazard* atau $H(t)$ dengan fungsi *survival* atau $S(t)$ sebagai berikut,

$$H(t) = -\ln S(t) \quad (4.7)$$

Jika resiko gagal (*failure*) pada waktu tertentu bergantung pada nilai x_1, x_2, \dots, x_p dari p variabel prediktor X_1, X_2, \dots, X_p maka nilai variabel tersebut diasumsikan telah tercatat sebagai *time origin*. Kumpulan nilai variabel prediktor dalam model *hazard proporsional* diwakili oleh vektor \mathbf{x} dengan $\mathbf{x} = (x_1, x_2, \dots, x_p)^T$. Misalkan $h_0(t)$ sebagai fungsi *hazard* untuk setiap obyek dengan nilai dari semua variabel prediktor penyusun vektor \mathbf{x} adalah nol maka fungsi $h_0(t)$ dikatakan sebagai fungsi *baseline hazard* (Collet, 2003). Model *hazard proporsional* dinyatakan seperti pada persamaan (2.8), namun untuk mengingatkan dituliskan kembali ke dalam nomor persamaan baru,

$$h(t, \mathbf{x}) = h_0(t) \exp(\mathbf{x}^T \boldsymbol{\beta}) \quad (4.8)$$

4.1.2. Penambahan efek random spasial (*spatial frailty term*) pada model *hazard proporsional*

Pada bab 2 telah dijelaskan bagaimana penambahan efek spasial dan dengan disertakan efek random pada model proporsional *hazard* yang tertulis pada persamaan (2.14), sedangkan untuk model *survival* dengan penambahan efek spasial pada persamaan disertakan efek random (2.16) dan telah dijelaskan juga pada bab 2 bahwa W_j pada persamaan (2.14) dan (2.16) merupakan bentuk *stratum-specific frailty* yang dibentuk untuk menyatakan perbedaan antar strata/kelompok. Secara umum W_j diasumsikan identik, independen dan

berdistribusi normal dengan mean bernilai nol dan varian σ^2 , akan tetapi menurut Hougaard (2000) dalam Banerjee *et al.*, (2003), distribusi non-normal juga sering digunakan untuk memodelkan *frailties*. Artinya tidak harus distribusi normal untuk asumsi W_j .

Model pada pendekatan *survival* spasial dibentuk dari data *survival* yang tersusun secara spasial, artinya *frailties* W_j dari kelompok/strata yang saling berdekatan menggambarkan kemungkinan bahwa kelompok/strata tersebut memiliki nilai/karakteristik yang serupa. Pendekatan lattice digunakan untuk menyusun susunan spasial dari kelompok-kelompok tersebut. Sehingga W_j terganti oleh W_j^* , hal tersebut dikarenakan \mathbf{W}^* didefinisikan oleh daerah-daerah dengan indeks diskret dimana daerah-daerah tersebut merupakan partisi/bagian dari daerah D . Sehingga model *hazard* proporsional yang ditambahkan efek spasial menjadi sebagai berikut :

$$h(t_{ij}; \mathbf{x}_{ij}) = h_0(t_{ij}) \exp(\beta_1 x_{1ij} + \beta_2 x_{2ij} + \dots + \beta_p x_{pij} + W_j^*) \quad (4.9)$$

atau dalam bentuk lainnya

model *hazard* proporsional unit i daerah ke-1

$$h(t_{i1}; \mathbf{x}_{i1}) = h_0(t_{i1}) \exp(\beta_1 x_{1i1} + \beta_2 x_{2i1} + \dots + \beta_p x_{pi1} + W_1^*)$$

model *hazard* proporsional unit i daerah ke-2

$$h(t_{i2}; \mathbf{x}_{i2}) = h_0(t_{i2}) \exp(\beta_1 x_{1i2} + \beta_2 x_{2i2} + \dots + \beta_p x_{pi2} + W_2^*)$$

\vdots

model *hazard* proporsional unit i daerah ke- j

$$h(t_{ij}; \mathbf{x}_{ij}) = h_0(t_{ij}) \exp(\beta_1 x_{1ij} + \beta_2 x_{2ij} + \dots + \beta_p x_{pij} + W_j^*)$$

\vdots

model *hazard* proporsional unit i daerah ke- m

$$h(t_{im}; \mathbf{x}_{im}) = h_0(t_{im}) \exp(\beta_1 x_{1im} + \beta_2 x_{2im} + \dots + \beta_p x_{pim} + W_m^*) \quad (4.10)$$

dengan : x_1, x_2, \dots, x_p adalah variabel prediktor ke-1 hingga ke- p

i menyatakan unit pengamatan, dimana unit pengamatan setiap lokasi berbeda jumlahnya.

j menyatakan daerah lokasi hingga lokasi ke- m

dikarenakan dalam penelitian ini terdapat 8 variabel prediktor dan 38 daerah kabupaten/kota, maka model umum *hazard* proporsional yang ditambahkan efek spasial menjadi sebagai berikut :

$$\begin{aligned} h(t_{i1}; \mathbf{x}_{i1}) &= h_0(t_{i1}) \exp(\beta_1 x_{1i1} + \beta_2 x_{2i1} + \dots + \beta_8 x_{8i1} + W_1^*) \\ &\vdots \\ h(t_{i38}; \mathbf{x}_{i38}) &= h_0(t_{i38}) \exp(\beta_1 x_{1i38} + \beta_2 x_{2i38} + \dots + \beta_8 x_{8i38} + W_{38}^*) \end{aligned} \quad (4.11)$$

4.1.3. Distribusi *prior* dan *joint* distribusi *posterior* dari model *mixture survival* spasial

Distribusi CAR (*Conditionally Autoregressive*) digunakan sebagai distribusi prior untuk parameter ν yang mewakili efek random (*frailty*) yang saling berkorelasi secara spasial, dapat dinyatakan sebagai berikut,

$$W^* | \lambda \sim \text{CAR}(\lambda) \quad (4.12)$$

dimana $\lambda = 1/\nu$ sedangkan bentuk umum dari prior CAR yang memiliki join distribusi proporsional sebagai berikut,

$$\lambda^{1/2} \exp \left[-\frac{\lambda}{2} \sum_{j \text{ adj } j'} (W_j^* - W_{j'}^*)^2 \right] \propto \lambda^{1/2} \exp \left[-\frac{\lambda}{2} \sum_{i=1}^I m_j W_j^* (W_j^* - \bar{W}_j^*) \right] \quad (4.13)$$

dimana : j dan j' adalah wilayah yang saling bertetangga

$$\bar{W}_j^* = \frac{\sum_{j'=1}^I W_{j'}^*}{m_j} \text{ yang merupakan rata-rata dari } W_{j'}^* \text{ yang bertetanggan dengan } W_j^*$$

m_j adalah jumlah tetangga yang dimiliki oleh daerah j

λ juga disebut sebagai parameter penghalus yang secara implisit juga menentukan variabilitas dari efek random (*frailty*). Nilai λ akan berbeda antara daerah satu dengan daerah yang lain, hal ini bergantung pada jumlah tetangga (kedekatan dengan daerah lain) yang dimiliki oleh daerah tersebut (m_j).

Pada penelitian ini, distribusi dari angka lama sekolah anak umur 16-18 tahun mengikuti sebenarnya berdistribusi nonnegatif distribusi karena berbentuk skew kiri, akan tetapi dalam pendekatan untuk setiap komponen *mixture* dari angka lama sekolah anak umur 16-18 tahun dilakukan dengan menggunakan

distribusi Normal 2-parameter (μ, σ) (penjelasan lebih lengkap pada subsub bab 4.2.5) sehingga fungsi kepadatan peluang (PDF) dari normal sebagai berikut,

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) \quad (4.14)$$

sedangkan fungsi distribusi kumulatifnya (CDF) sebagai berikut,

$$\begin{aligned} F(t) = P(T \leq t) &= \int_{-\infty}^t \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) dt \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^t \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) d(t) \end{aligned}$$

misalkan

$$\begin{aligned} \frac{t-\mu}{\sigma} &= x \\ t &= x\sigma + \mu \\ dt &= \sigma dx \end{aligned}$$

maka

$$\begin{aligned} F(t) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^t \exp\left(-\frac{1}{2}(x)^2\right) \sigma dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 \exp\left(-\frac{1}{2}x^2\right) d(x) + \frac{1}{\sqrt{2\pi}} \int_0^t \exp\left(-\frac{1}{2}x^2\right) d(x) \\ &= \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{t-\mu}{\sqrt{2}\sigma}\right) \end{aligned} \quad (4.15)$$

Nilai t selalu bernilai positif dikarenakan waktu survival angka lama dari individu untuk bertahan dari suatu kondisi maka selalu bernilai positif. Pada hasil CDF pada 4.15 terdapat hasil erf dimana erf merupakan *error function* dari CDF

distribusi Normal dikarenakan $\frac{1}{\sqrt{2\pi}} \int_0^t \exp\left(-\frac{1}{2}x^2\right) d(x)$ memiliki hasil yang tidak *close form* sehingga terbentuklah *error function*.

Berdasarkan persamaan CDF pada (4.15) dan fungsi *survival* pada persamaan (2.3), maka dapat ditentukan fungsi *survival* dari distribusi Normal 2-parameter sebagai berikut,

$$S(t) = 1 - F(t) = 1 - \left(\frac{1}{2} + \frac{1}{2} \operatorname{erf} \left(\frac{t - \mu}{\sqrt{2\sigma^2}} \right) \right) = \frac{1}{2} - \frac{1}{2} \operatorname{erf} \left(\frac{t - \mu}{\sqrt{2\sigma^2}} \right) \quad (4.16)$$

Untuk menentukan fungsi *hazard* dari distribusi Normal 2-parameter maka diperlukan PDF pada persamaan (4.14) dan fungsi *survival* pada persamaan (4.16), selanjutnya kedua persamaan tersebut dioperasikan melalui persamaan (4.3) sebagai berikut,

$$h(t) = \frac{f(t)}{S(t)} = \frac{\frac{1}{\sigma\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{t - \mu}{\sigma} \right)^2 \right)}{\frac{1}{2} - \frac{1}{2} \operatorname{erf} \left(\frac{t - \mu}{\sqrt{2\sigma^2}} \right)} = \frac{\sqrt{2} \exp \left(-\frac{1}{2} \left(\frac{t - \mu}{\sigma} \right)^2 \right)}{\sqrt{\pi\sigma^2} \left(1 - \operatorname{erf} \left(\frac{t - \mu}{\sqrt{2\sigma^2}} \right) \right)} \quad (4.17)$$

Berdasarkan fungsi *hazard* pada persamaan (4.17) dan (2.9) maka dapat diperoleh persamaan berikut,

$$h(t, \mathbf{x}) = h_0(t) \exp(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p) = \frac{\sqrt{2} \exp \left(-\frac{1}{2} \left(\frac{t - \mu}{\sigma} \right)^2 \right)}{\sqrt{\pi\sigma^2} \left(1 - \operatorname{erf} \left(\frac{t - \mu}{\sqrt{2\sigma^2}} \right) \right)} \quad (4.18)$$

Pada persamaan (4.18), dikarenakan masih terdapat *error function*, cukup susah untuk menentukan mana yang merupakan bagian dari $h_0(t)$ tetapi untuk $\exp(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)$ dituliskan sebagai berikut :

$$\exp(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p) = \mu \quad (4.19)$$

Sebelum melakukan estimasi terhadap parameter μ dan σ (pada syntax simbol μ disebut psi dan simbol σ disebut tau), maka terlebih dahulu menentukan distribusi priornya sebagai berikut,

$$\sigma \sim \text{Gamma}(p, q)$$

$$\beta_i \sim \text{Normal}(m, s)$$

dimana penentuan dimana penentuan distribusi prior tersebut dilakukan berdasarkan gabungan antara *prior conjugate* dan *prior informatif* yaitu *prior* yang ditentukan berdasarkan pola *likelihood* data dan berdasarkan informasi *prior* dari penelitian sebelumnya. Pada analisis *survival* menyertakan data tersensor dalam menentukan join distribusi posteriornya, γ_{ij} menyatakan status tersensor dari responden angka lama sekolah yaitu nilai 0 untuk responden tersensor dan

nilai 1 untuk responden tidak tersensor. Selanjutnya ditentukan t sebagai waktu *survival* responden (waktu yang diperlukan hingga terjadinya lulus) dan x merupakan vektor *covariate* (variabel prediktor) sehingga diperoleh join distribusi posteriornya sebagai berikut,

$$P(\beta, W, \sigma, \lambda | t, x, \gamma) \propto L(\beta, W, \sigma; t, x, \gamma) p(W | \lambda) p(\beta) p(\sigma) p(\lambda) \quad (4.20)$$

Pada ruas kanan terdapat lima bentuk, yaitu bentuk pertama merupakan likelihood untuk *hazard* distribusi Normal 2-parameter, bentuk kedua menyatakan join distribusi dari efek random (*frailty*), sedangkan tiga bentuk sisanya merupakan distribusi prior dari masing-masing parameter.

Fungsi likelihood untuk *hazard* distribusi Normal 2-parameter dapat dijabarkan sebagai berikut (Collet, 2003),

$$L(\beta, W, \sigma; t, x, \gamma) = \prod_{i=1}^I \prod_{j=1}^{n_i} f_{ij}(t_{ij})^{\gamma_{ij}} S_{ij}(t_{ij})^{1-\gamma_{ij}} \quad (4.21)$$

Dengan menggunakan persamaan (4.4) sehingga dapat ditulis $f(t) = h(t)S(t)$. Selanjutnya persamaan tersebut dapat disubstitusikan ke persamaan berikut,

$$\begin{aligned} L(\beta, W, \sigma; t, x, \gamma) &= \prod_{i=1}^I \prod_{j=1}^{n_i} f_{ij}(t_{ij})^{\gamma_{ij}} S_{ij}(t_{ij})^{1-\gamma_{ij}} \\ &= \prod_{i=1}^I \prod_{j=1}^{n_i} [h(t_{ij}) S_{ij}(t_{ij})]^{\gamma_{ij}} S_{ij}(t_{ij})^{1-\gamma_{ij}} \\ &= \prod_{i=1}^I \prod_{j=1}^{n_i} h(t_{ij})^{\gamma_{ij}} S_{ij}(t_{ij})^{\gamma_{ij}} S_{ij}(t_{ij})^{1-\gamma_{ij}} \\ &= \prod_{i=1}^I \prod_{j=1}^{n_i} h(t_{ij})^{\gamma_{ij}} S_{ij}(t_{ij}) \end{aligned} \quad (4.22)$$

Sedangkan untuk parameter σ , β_i , dan λ memiliki distribusi posterior marginal dengan cara mengintergralkan keluar parameter-parameter yang bersangkutan dapat dijabarkan sebagai berikut,

$$p(\sigma | \lambda, \beta_i) \propto \int_{\lambda} \int_{\beta_1} \cdots \int_{\beta_p} l(t | \lambda, \beta_1, \dots, \beta_p) p(\lambda) p(\beta_1) \dots p(\beta_p) d\lambda d\beta_1, \dots, d\beta_p$$

$$p(\lambda | \sigma, \beta_i) \propto \int_{\sigma} \int_{\beta_1} \cdots \int_{\beta_p} l(t | \sigma, \beta_1, \dots, \beta_p) p(\sigma) p(\beta_1) \dots p(\beta_p) d\sigma d\beta_1, \dots, d\beta_p$$

$$\begin{aligned} p(\beta_1 | \sigma, \lambda, \beta_{i \neq 1}) &\propto \int_{\sigma} \int_{\beta_2} \cdots \int_{\beta_p} l(t | \sigma, \lambda, \beta_2, \dots, \beta_p) p(\sigma) p(\lambda) \\ &\quad \times p(\beta_2) \dots p(\beta_p) d\sigma d\lambda d\beta_2, \dots, d\beta_p \end{aligned}$$

$$p(\beta_p | \sigma, \lambda, \beta_{i \neq p}) \propto \int \int \cdots \int_a \int_{\beta_2} \int_{\beta_p} l(t | \sigma, \lambda, \beta_1, \dots, \beta_{p-1}) p(\sigma) p(\lambda) \\ \times p(\beta_1) \dots p(\beta_{p-1}) d\sigma d\lambda d\beta_1, \dots, d\beta_{p-1}$$

4.1.4. Estimasi Parameter Model Mixture Survival Spasial Dengan Menggunakan RJMCMC dan Gibbs Sampling

Pada bab 3 di penelitian ini sudah dijelaskan bahwa dilakukan pengujian dan penentuan distribusi dari data angka lama sekolah anak umu 16-18 tahun yang akan digunakan untuk model *survival* dan model *hazard* nya, dan jika terdapat indikasi bahwa distribusi dari data angka lama sekolah anak umur 16-18 tahun terdapat *mixture* dan cukup sulit untuk menentukan komponen *mixture*nya maka akan digunakan algoritma *Reversible jump Markov Chain Monte Carlo* (RJMCMC).

Keuntungan dari penggunaan metode ini adalah dapat memodelkan suatu *mixture* secara bersama-sama dengan banyaknya komponen penyusun *mixture* dan algoritma ini akan memudahkan untuk mendapatkan model *mixture* yang tepat bagi data yang terindikasi bersifat multimodal. Algoritma ini memanfaatkan proses *birth/death* dan *split/merge* dalam aplikasinya seperti yang tertulis pada algoritma pada Bab 2. Sedangkan untuk menentukan parameter pada model dengan menggunakan MCMC dengan *Gibbs Sampling*.

4.2. Model Mixture Survival Spasial Pada Angka Lama Sekolah Anak Umur 16-18 Tahun di Jawa Timur Berdasarkan Variabel-Variabel Yang Mempengaruhi.

Dalam sub bab ini akan dicari model *mixture survival* spasial pada angka lama sekolah anak umur 16-18 tahun di Jawa Timur berdasarkan variabel-variabel yang mempengaruhi tinggi atau rendahnya angka lama sekolah anak umur 16-18 tahun. Namun sebelumnya akan dibahas karakteristik responden anak umur 16-18 tahun berdasarkan variabel-variabel prediktornya.

4.2.1. Karakteristik Responden Anak Umur 16-18 Tahun Berdasarkan Variabel-Variabel Prediktornya

Pada sub bab ini dilakukan analisis statistika deskriptif untuk mengetahui karakteristik responden anak yang berumur 16-18 tahun. Analisis deskriptif ini dilakukan terhadap masing-masing variabel dalam penelitian meliputi variabel respon yaitu angka lama sekolah dan 8 variabel prediktor antara lain, variabel jenis kelamin (X_1), status perkawinan (X_2), status bekerja (X_3), tingkat pendidikan kepala rumah tangga (X_4), jenis kelamin kepala rumah tangga (X_5), banyaknya anggota rumah tangga (X_6), rata-rata pengeluaran rumah tangga per kapita (X_7), dan lokasi tempat tinggal (X_8). Untuk variabel yang berskala rasio akan dilihat nilai *mean*, varian, minimum dan maksimum dari variabel untuk mendapatkan informasi tentang variabel tersebut, sedangkan untuk variabel yang berskala nominal maupun ordinal untuk mendapatkan informasi tentang variabel tersebut akan dilihat dalam bentuk diagram batang dan *pie chart*. Data angka lama sekolah anak umur 16-18 tahun yang diolah dapat dilihat pada Lampiran 1.

Tabel 4.1. Deskriptif Responden Anak Berumur 16-18 Tahun

Variabel	Minimum	Maksimum	Mean	Std. Deviation
Angka Lama Sekolah Anak Umur 16-18 Tahun	0	12	9.39	1.841
Jumlah Anggota Rumah Tangga	1	11	4.41	1.342
Rata-Rata Pengeluaran Rumah Tangga	120955,94	13113030,12	517887,26	481380,3477

Sumber : pengolahan SPSS

Pada Tabel 4.1 didapatkan informasi bahwa waktu *survival* responden atau angka lama sekolah anak umur 16-18 tahun paling kecil 0 tahun yang artinya belum pernah menyentuh pendidikan formal. Paling besar yaitu 12 tahun yang artinya sudah mendapatkan ijazah pada jenjang SMA/ sederajat, sedangkan mean atau rata-rata nya 9,39 yang artinya sebagian besar anak yang berumur 16-18 tahun memiliki rata-rata lama sekolah di atas 9 tahun atau sudah lulus pada jenjang SMP/ sederajat. Untuk variabel jumlah anggota keluarga, paling sedikit ada rumah tangga yang memiliki 1 anggota keluarga dan yang paling banyak

hingga 11 anggota keluarga di dalamnya. Sedangkan untuk variabel rata-rata pengeluaran rumah tangga per kapita per bulan, pengeluaran terendah yaitu Rp.120.955,94 dan pengeluaran tertinggi hingga Rp.13.113.030,12, sedangkan untuk rata-rata dari rata-rata pengeluaran rumah tangga yaitu Rp.517.887,26.

Tabel 4.2. Deskriptif Status Sensor Anak Umur 16-18 Tahun Per Kabupaten/Kota Se Jawa Timur.

Kabupaten/Kota	Tersensor	%	Tidak Tersensor	%	Total
Pacitan	15	19%	66	81%	81
Ponorogo	10	12%	76	88%	86
Trenggalek	21	24%	66	76%	87
Tulungagung	16	14%	98	86%	114
Blitar	15	19%	62	81%	77
Kediri	28	23%	94	77%	122
Malang	34	25%	103	75%	137
Lumajang	24	29%	60	71%	84
Jember	30	24%	97	76%	127
Banyuwangi	22	18%	97	82%	119
Bondowoso	25	32%	54	68%	79
Situbondo	19	26%	54	74%	73
Probolinggo	23	29%	55	71%	78
Pasuruan	39	30%	89	70%	128
Sidoarjo	8	6%	120	94%	128
Mojokerto	14	9%	138	91%	152
Jombang	16	13%	110	87%	126
Nganjuk	14	14%	86	86%	100
Madiun	12	13%	82	87%	94
Magetan	8	12%	57	88%	65
Ngawi	8	9%	81	91%	89
Bojonegoro	21	17%	105	83%	126
Tuban	14	13%	97	87%	111
Lamongan	7	8%	86	92%	93
Gresik	8	6%	128	94%	136
Bangkalan	71	53%	63	47%	134
Sampang	60	58%	43	42%	103
Pamekasan	30	26%	86	74%	116
Sumenep	23	27%	63	73%	86
Kota Kediri	9	11%	75	89%	84
Kota Blitar	4	7%	55	93%	59
Kota Malang	13	13%	86	87%	99

Sumber : Pengolahan SPSS

Lanjutan Tabel 4.2

Kabupaten/Kota	Tersensor	%	Tidak Tersensor	%	Total
Kota Probolinggo	9	11%	73	89%	82
Kota Pasuruan	9	11%	73	89%	82
Kota Mojokerto	6	9%	61	91%	67
Kota Madiun	9	14%	54	86%	63
Kota Surabaya	19	12%	136	88%	155
Kota Batu	5	10%	47	90%	52
Jawa Timur	718	19%	3076	81%	3794

Deskriptif untuk angka lama sekolah anak umur 16-18 tahun di Jawa Timur dengan pengkategorian menjadi tersensor dan tidak tersensor dapat dilihat pada Tabel 4.2. bahwa untuk kategori tersensor dengan angka lama sekolah < 9 tahun, sedangkan untuk kategori tidak tersensor dengan rentang angka lama sekolah 9-12 tahun, maka didapatkan informasi bahwa jumlah anak yang memiliki angka lama sekolah 9-12 tahun atau dikatakan tidak tersensor pada Susenas 2012 sebanyak 3076 anak atau 81% dari jumlah anak yang berumur 16-18 tahun yang dipilih sebagai responden dalam Susenas 2012 sedangkan yang tidak tersensor dengan angka lama sekolah < 9 tahun atau dikatakan tidak lulus SMP/ sederajat sesuai program pemerintah dalam hal ini disebut tersensor sebanyak 718 anak atau 19% dari jumlah anak yang berumur 16-18 tahun yang dipilih sebagai responden dalam Susenas 2012. Angka lama sekolah anak umur 16-18 tahun di Jawa Timur berdasarkan kabupaten/kota didapatkan jumlah rumah tangga yang memiliki anak dengan umur 16-18 tahun disetiap kabupaten/kota beragam jumlahnya. Kabupaten/kota dengan jumlah rumah tangga yang terpilih sebagai responden dan memiliki anak umur 16-18 paling banyak yaitu kota Surabaya dengan 155 rumah tangga, sedangkan yang paling sedikit yaitu kota Batu dengan 52 rumah tangga.

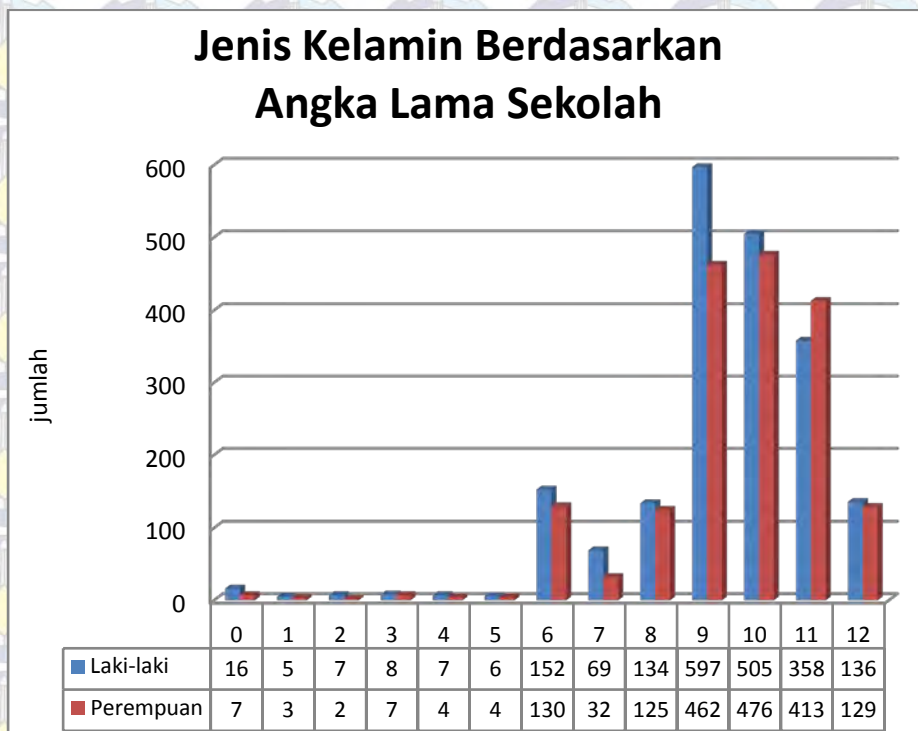
Jika melihat persentase dari jumlah anak yang tersensor dan yang tidak tersensor, yang tersensor paling banyak kabupaten Sampang dengan 60 anak atau 58% dari jumlah anak yang berumur 16-18 tahun yang dipilih sebagai responden di kabupaten Sampang tidak lulus pada jenjang SMP/ sederajat dan yang tersedikit kota Batu dengan 5 anak tetapi jika dilihat persentase yang tersedikit yaitu kabupaten Sidoarjo dengan 8 anak atau 6% dari jumlah anak yang berumur 16-18 tahun yang dipilih sebagai responden di kabupaten Sidoarjo tidak lulus pada

jenjang SMP/ sederajat. Sedangkan untuk tidak tersensor, paling banyak secara persentase yaitu kabupaten Gresik dengan 94 % (128 anak) dari jumlah anak yang berumur 16-18 tahun yang dipilih sebagai responden di kabupaten Gresik lulus pada jenjang SMP/ sederajat. Untuk yang tersedikit secara persentase yaitu kabupaten Sampang dengan 42% (43 anak) dari jumlah anak yang berumur 16-18 tahun yang dipilih sebagai responden di kabupaten Sampang lulus pada jenjang SMP/ sederajat.

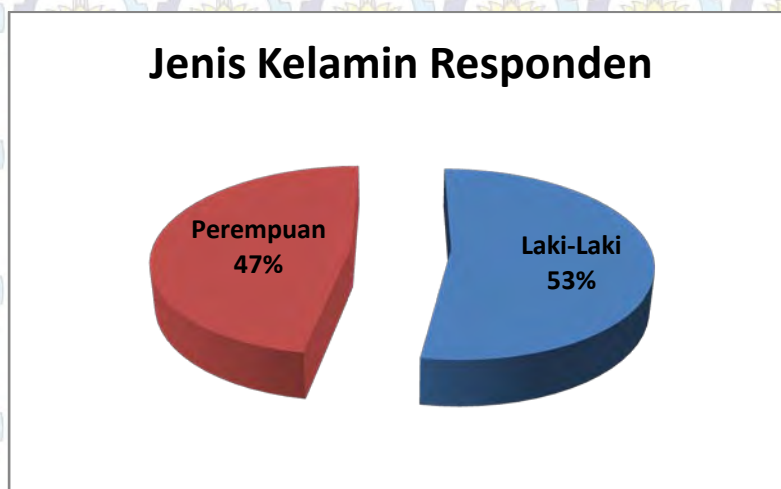
Seperti yang telah dijelaskan pada awal subbab ini, deskriptif dari variabel prediktor yang berkategori nominal dan ordinal akan dijelaskan dengan secara visual dengan menggunakan digram batang dan pie chart. Variabel-variabel yang akan dijelaskan dengan visual adalah variabel jenis kelamin responden, status perkawinan responden, status bekerja responden, jenis kelamin kepala rumah tangga, tingkat pendidikan terakhir kepala rumah tangga, dan lokasi tempat tinggal responden.

Jumlah responden anak laki-laki umur 16-18 tahun lebih banyak dari responden perempuan, hal tersebut dapat dilihat pada Gambar 4.2, jumlah responden laki-laki mencapai 53% dari jumlah responden anak umur 16-18 tahun secara keseluruhan sedangkan untuk responden perempuan hanya 47%. Sedangkan untuk perbedaan jumlah responden laki-laki dan perempuan disetiap angka lama sekolah dapat dilihat pada Gambar 4.1, bahwa disetiap jenjang angka lama sekolah responden laki-laki lebih banyak dari responden perempuan, kecuali pada angka lama sekolah 11 dimana jumlah responden perempuan lebih banyak.

Pada Gambar 4.1 juga dapat dilihat bahwa jumlah responden cenderung lebih banyak pada jenjang angka lama sekolah 9-12 tahun baik untuk responden laki-laki maupun responden perempuan.

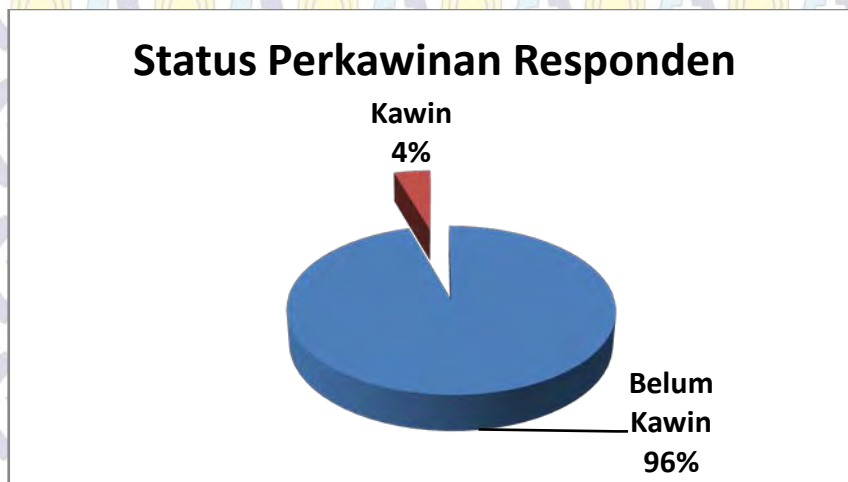


Gambar 4.1. Diagram Batang Variabel Jenis Kelamin Berdasarkan Angka Lama Sekolah Anak Umur 16-18 Tahun

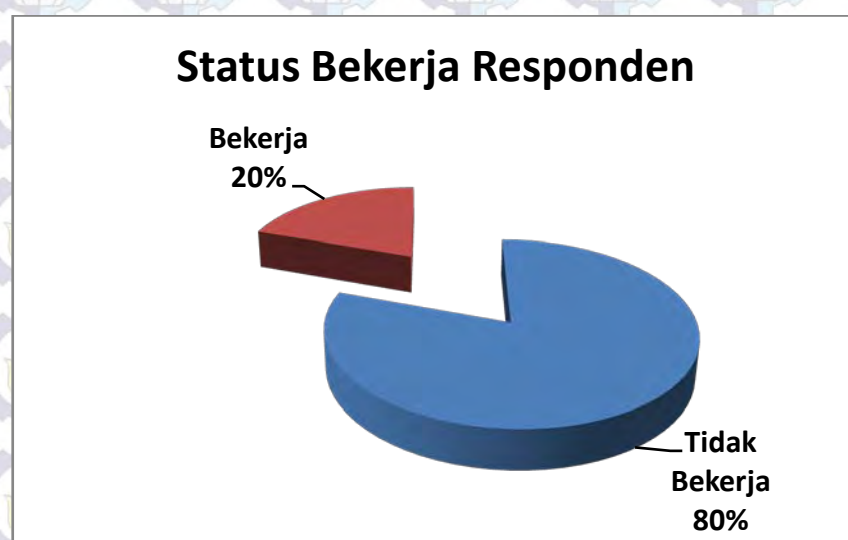


Gambar 4.2. Pie Chart Jenis Kelamin Responden

Pada variabel status perkawinan dari responden, terdapat 4% (sekitar 152 anak) responden yang sudah kawin dan sebagian besar dengan 96% belum kawin. Untuk status bekerja responden terdapat 20% responden yang sudah bekerja pada umur 16-18 tahun sedangkan sisanya 80% tidak bekerja.



Gambar 4.3. Pie Chart Status Perkawinan dari Responden



Gambar 4.4. Pie Chart Status Bekerja dari Responden

Berdasarkan Tabel 4.3. dapat dilihat bahwa responden anak umur 16-18 tahun dengan status bekerja lebih banyak pada rumah tangga dengan rata-rata pengeluaran <Rp. 500.000 yaitu 594 anak, sedangkan untuk anak dengan penghasilan rumah tangga pada rentang Rp. 500.001 – Rp. 1.000.000 hanya 130 anak dan sisanya 18 anak dengan status bekerja pada rumah tangga dengan pengeluaran > Rp. 1.000.000. Untuk status responden anak yang tidak bekerja lebih banyak dari yang bekerja, hal tersebut mungkin dikarenakan orangtua

mereka tidak memperkenankan anak mereka untuk bekerja dahulu pada usia sekolah.

Tabel 4.3. Deskriptif Rata-Rata Pengeluaran terhadap Status Bekerja Responden

Rata-Rata Pengeluaran	Status_Bekerja		Total
	Tidak	Ya	
<Rp. 500.000	1887	594	2481
Rp. 500.001 - Rp. 1000.000	912	130	1042
>Rp. 1000.001	253	18	271
Total	3052	742	3794

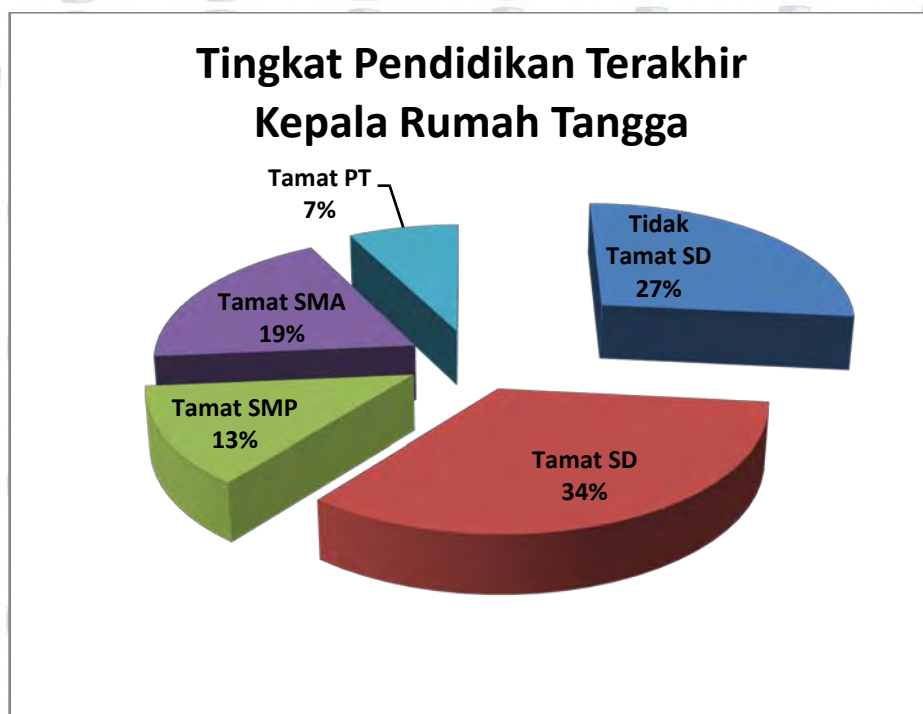
Dari penelitian sebelumnya tentang angka lama sekolah yang terdahulu, jenis kelamin kepala rumah tangga mempengaruhi angka lama sekolah dari anak-anak mereka. Pada penelitian ini kepala rumah tangga berjenis kelamin laki-laki lebih banyak dari kepala rumah tangga berjenis kelamin perempuan, yaitu 90% (3410 kepala rumah tangga) yang berjenis kelamin laki-laki dan 10% (384 kepala rumah tangga) yang berjenis kelamin perempuan.



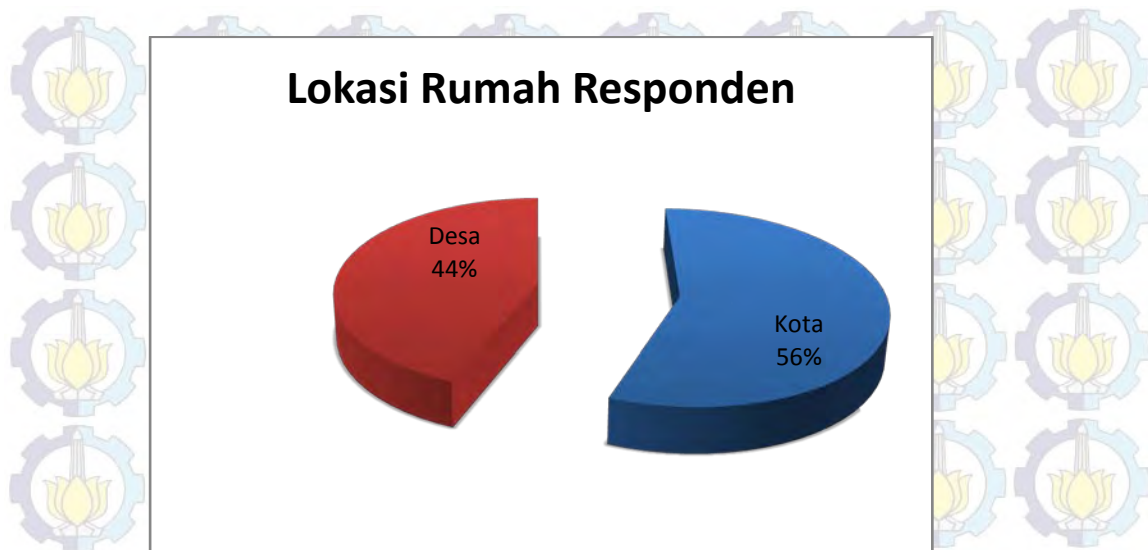
Gambar 4.5. Pie Chart Jenis Kelamin Kepala Rumah Tangga Responden

Sedangkan untuk tingkat pendidikan terakhir kepala rumah tangga, bahwa 27 % (1009 kepala rumah tangga) tidak tamat SD/ sederajat atau dapat dikatakan tidak mendapat pendidikan formal, 34% (1293 kepala rumah tangga) hanya memiliki ijazah sampai jenjang SD/ sederajat, 13% (490 kepala rumah tangga) memiliki ijazah sampai jenjang SMP/ sederajat, 19% (714 kepala rumah tangga) memiliki ijazah sampai jenjang SMA/ sederajat dan sisanya 7% (288 kepala rumah tangga) memiliki ijazah sampai perguruan tinggi.

Untuk variabel lokasi tempat tinggal responden, 56% dari total responden atau 2133 responden yang terpilih bertempat tinggal di perkotaan, sedangkan sisanya 44% dari total responden atau 1661 responden yang terpilih bertempat tinggal di perdesaan.



Gambar 4.6. Pie Chart Tingkat Pendidikan Terakhir Kepala Rumah Tangga Responden

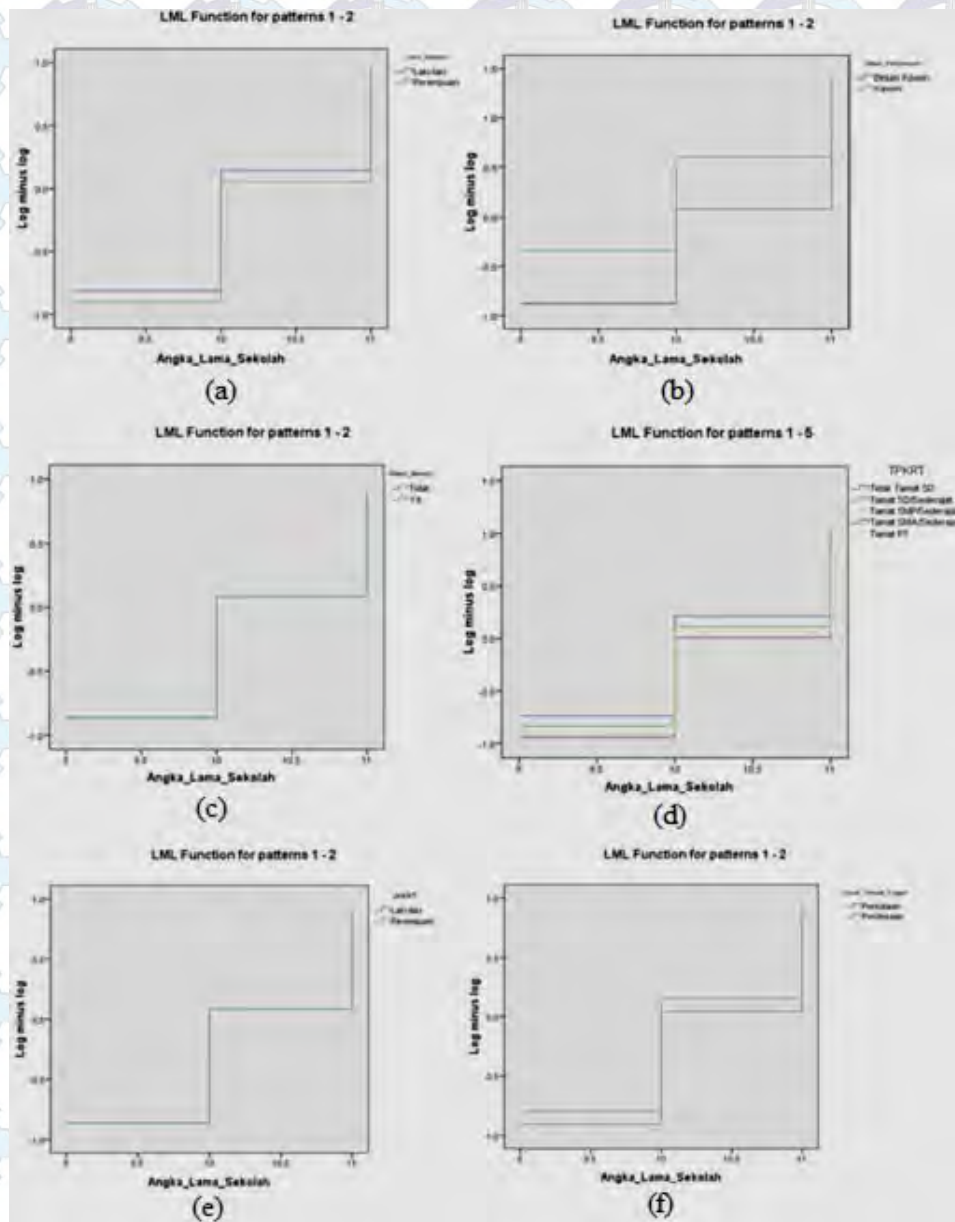


Gambar 4.7. Pie Chart Lokasi Rumah Responden

4.2.2. Pengujian Asumsi Pemodelan *Hazard Proporsional*

Asumsi pemodelan yang harus dipenuhi dalam *hazard* proporsional adalah bahwa fungsi *hazard* dari variabel prediktor yang bersifat kategorik harus proporsional setiap waktu. Pengujian dapat dilakukan dengan menggunakan plot $\ln[-\ln S(t)]$. Asumsi *hazard* proporsional terpenuhi bila garis antar kategori sejajar.

Telah dijelaskan pada bab II bahwa pada penelitian ini terdapat enam variabel yang bersifat kategorik dari delapan variabel. Dari Gambar 4.8 (a), (b), (c), (d), (e), dan (f), dapat dilihat bahwa garis antar kategori sejajar sehingga bisa dikatakan asumsi *hazard* proporsional sudah terpenuhi untuk keenam variabel kategorik, artinya variabel prediktor telah independen terhadap waktu serta hubungan antara kumulatif *hazard* sudah proporsional/ konstan setiap waktu.



Gambar 4.8. Plots $\ln[-\ln S(t)]$ Untuk Uji Proporsional Hazard (a). Jenis Kelamin Responden, (b). Status Perkawinan Responden, (c). Status Bekerja Responden, (d). Tingkat Pendidikan Terakhir Kepala Rumah Tangga, (e). Jenis Kelamin Kepala Rumah Tangga, dan (f). Lokasi Tempat Tinggal Responden.

4.2.3. Penimbang Spasial

Pendekatan yang digunakan dalam menentukan penimbang spasial dalam penelitian ini menggunakan pendekatan area. Penimbang spasial yang diperoleh selanjutnya merupakan salah satu parameter dalam prior CAR yang menjadi distribusi dari efek random (*frailty*) model *survival*. Jenis penimbang yang

digunakan yaitu penimbang spasial *Queen Contiguity* (persinggungan sisi dan sudut) dan kabupaten yang saling bertetangga disebut sebagai *adjacent* sehingga ketetanggaan dari masing-masing kabupaten di provinsi Jawa Timur akan disusun dalam matriks *adjacent*.

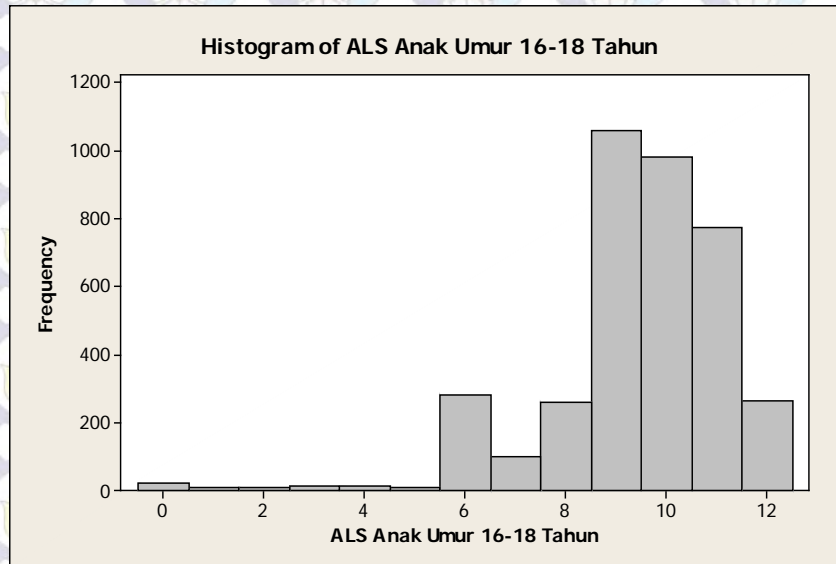
Langkah dalam menentukan matriks *adjacent* dengan mengolah peta/map poligon Jawa Timur per kabupaten yang sudah didapat melalui *software* GeoDa. Hasil dari penimbang spasial *Queen Contiguity* yang telah didapat akan berupa matriks dengan ukuran 38×38 (dikarenakan banyak kabupaten/kota di Jawa Timur sebanyak 38 kabupaten/kota) yang element dari matriks tersebut 0 dan 1, 1 untuk daerah yang bertetangga dengan daerah j (daerah yang dilihat ketetanggannya), sedangkan 0 untuk daerah yang tidak bertetangga dengan daerah j . Lalu matriks ini distandarisasi dengan cara membagi setiap barisnya dengan banyak daerah yang bertetangga dengan daerah j . Hasil matriks yang sudah distandarisasi dapat dilihat pada lampiran. Secara keseluruhan matriks *adjacent* ini memiliki 140 ketetanggaan dari 38 kabupaten yang diteliti.

4.2.4. Mengidentifikasi dan Menguji Distribusi Data Waktu *Survival* Angka Lama Sekolah

Pendugaan distribusi dilakukan terhadap waktu *survival* (t) dimana pada penelitian ini adalah angka lama sekolah anak umur 16-18 tahun di Jawa Timur. Gambar 4.9 menunjukkan histogram dari waktu *survival* responden anak umur 16-18 tahun, dapat dilihat bahwa diindikasikan terdapat lebih dari satu puncak sehingga diduga terdapat kasus multimodal pada waktu *survival* ini sehingga perlu digunakan distribusi *mixture* agar hasil estimasi distribusi dan hasil analisis selanjutnya bisa lebih akurat, hal tersebut didukung dengan hasil pengujian Anderson-Darling pada Lampiran 2 dengan tidak ditemukan distribusi yang sesuai untuk data angka lama sekolah anak umur 16-18 tahun.

Tidak semua distribusi *mixture* bisa terbaca atau dapat ditentukan berapa banyak komponennya hanya dengan gambaran visual dari histogram saja. Metode menentukan banyak komponen *mixture* dengan melihat visual dari histogram disebut metode prediktif, dimana hanya memprediksi berapa banyak komponen *mixture* berdasarkan banyak puncak dari data. Untuk itu diperlukan suatu metode dimana dimana penentuan jumlah komponen *mixture*-nya lebih akurat dan dapat

dipercaya, untuk itu dalam penelitian ini untuk penentuan banyak komponen distribusi *mixture* agar lebih akurat digunakan metode *reversible jump* markov chain monte carlo.



Gambar 4.9. Histogram Angka Lama Sekolah Anak Umur 16-18 Tahun

Dengan hipotesis bahwa

H0 : Waktu *survival* sesuai dengan distribusi dugaan (Lognormal/Normal/Weibull/Log Logistik)

H1 : Waktu *survival* tidak sesuai dengan distribusi dugaan (Lognormal/Normal/Weibull/Log Logistik)

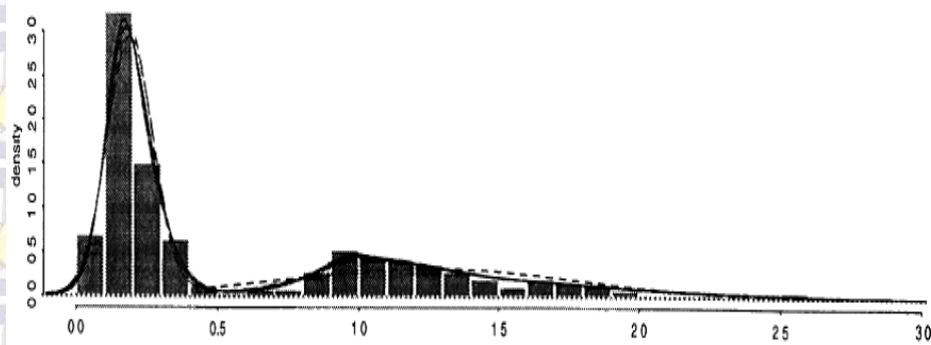
Statistik uji sesuai dengan persamaan (2.1) dan daerah kritis didasarkan pada kondisi tolak H₀ jika $A_n^2 > a_{n,1-\alpha}$ atau P-value $< \alpha$, dengan $\alpha = 0,05$ maka hasilnya dapat dilihat pada Tabel 4.4 dan Lampiran 2. Distribusi yang paling mendekati untuk data angka lama sekolah anak umur 16-18 tahun yaitu distribusi Normal 2-parameter. Distribusi Normal 2-parameter ini akan dijadikan distribusi utama dalam penerapan *reversible jump* markov chain monte carlo.

Tabel 4.4. Uji Distribusi Waktu *Survival*

Distribusi	Statistik Uji (A_n^2)	Nilai Kritis ($\alpha_{n,1-\alpha}$)	Keputusan
Log Normal	281,78	2,5018	Tolak H_0
Normal	146,06	2,5018	Tolak H_0
Weibull	147,00	2,5018	Tolak H_0
Log Logistik	193,50	2,5018	Tolak H_0

Sumber : Program Easy Fit

Distribusi Normal adalah salah satu distribusi non-negatif. sedangkan jika dilihat berdasarkan Gambar histogram (4.9), tidak ada angka yang bernilai negatif. Pendekatan dengan distribusi normal yang dimaksud untuk mendekati setiap bagian komponen *mixture* dengan distribusi normal.

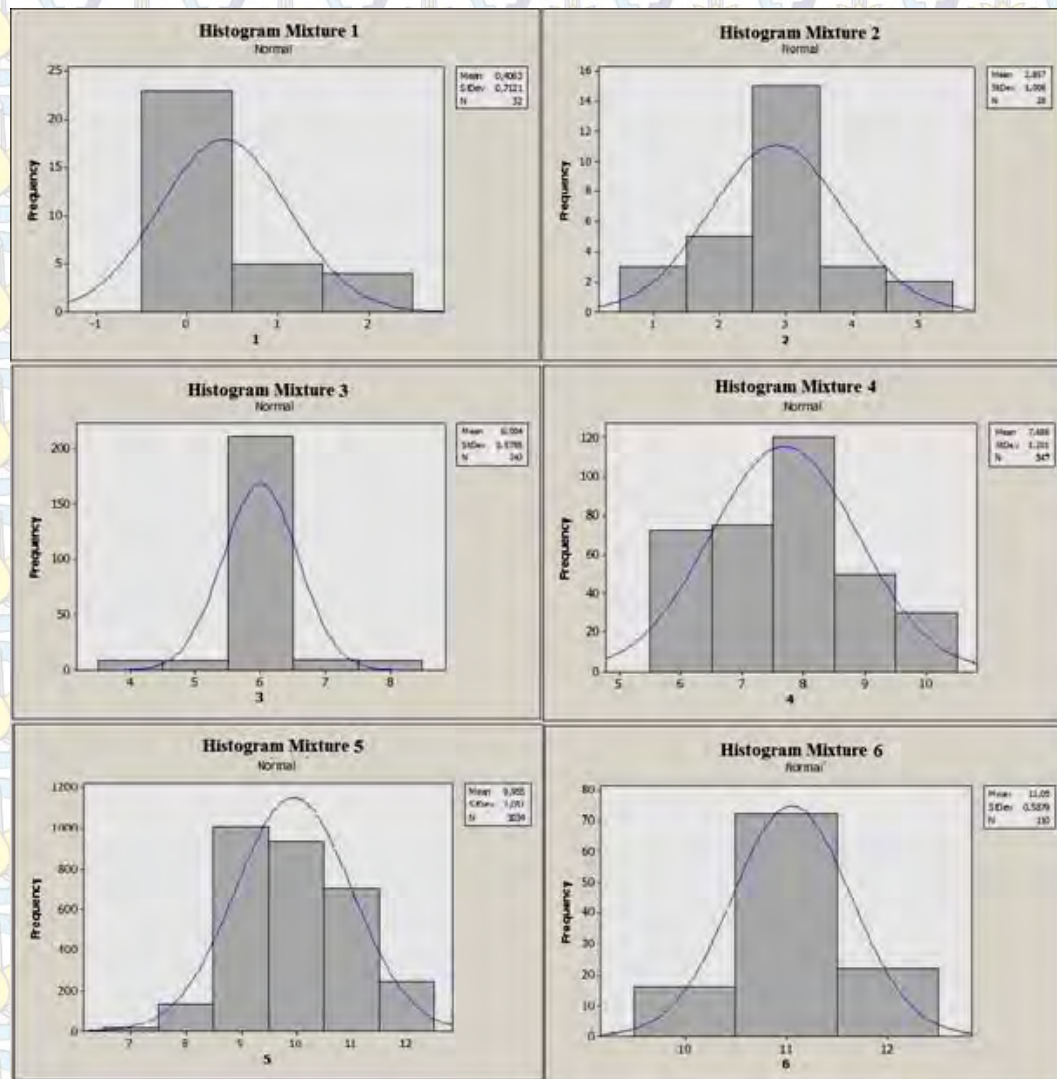


Gambar 4.10. Contoh Pendekatan Distribusi Normal Pada Komponen *Mixture* Pada Penelitian Green dan Sylvia (1997) Data Enzyme

4.2.5. Menentukan Model dan Parameter Model *Mixture Survival* Spasial

Sebelum menentukan model *mixture survival* spasial akan ditentukan dahulu banyak komponen yang dibentuk oleh model dengan menggunakan algoritma RJMCMC, adapun syntax dan hasil dapat dilihat pada Lampiran 5. Berdasarkan hasil yang diperoleh, dengan menggunakan iterasi sebanyak 100000, didapatkan banyak komponen untuk distribusi *mixture* pada data *survival* angka lama sekolah anak umur 16-18 tahun yaitu sebanyak 6 komponen (lihat Gambar 4.11). Hal tersebut dapat dilihat bahwa hanya 6 komponen yang memiliki nilai posterior probability yang menjadi nilai komponen awal setiap *mixture*, yaitu sebesar 0,86021, 0,05878, 0,05303, 0,01396, 0,00557, dan 0,00845. Nilai

komponen awal ini digunakan untuk menentukan data yang sesuai dengan komponennya.



Gambar 4.11. Histogram Data *Mixture* 1 Sampai *Mixture* 6

Jumlah komponen *mixture* yang telah diketahui digunakan untuk membentuk model *mixture survival* spasial dengan variabel-variabel yang mempengaruhi. Digunakan software WinBUGS untuk membantu dalam perhitungan serta penentuan nilai parameter-parameter pada persamaan 4.20.

Tabel 4.5. Estimasi Parameter Model *Mixture Survival Spasial*

Node	Mean	Sd	MC Error	2.5%	Median	97.5%
Phi[1]	0,00869	0,001481	0,0332	0,005987	0,0086	0,0118
Phi[2]	0,00767	0,001393	0,0279	0,005288	0,007587	0,01066
Phi[3]	0,0642	0,004079	0,0913	0,05663	0,06416	0,07238
Phi[4]	0,09174	0,004648	0,00008	0,08291	0,0916	0,101
Phi[5]	0,7984	0,006549	0,133	0,7856	0,7986	0,8113
Phi[6]	0,02929	0,002773	0,0561	0,02411	0,02919	0,03501
b1[1]	0,02705	0,08932	0,01272	-0,1669	0,02487	0,1964
b2[1]	0,9876	0,1508	0,02214	0,7003	0,9863	1,288
b3[1]	1,328	0,1621	0,02415	1,091	1,32	1,599
b4_1[1]	-0,08707	0,1801	0,02625	-0,5033	-0,07932	0,2023
b4_2[1]	0,07857	0,3273	0,04377	-0,6468	0,07265	0,7583
b4_3[1]	-0,1805	0,2108	0,03023	-0,5683	-0,1531	0,1717
b4_4[1]	-0,08752	0,3676	0,0499	-0,7417	-0,1116	0,7962
b5[1]	0,4103	0,09288	0,0131	0,2095	0,408	0,5642
b6[1]	0,003723	0,02156	0,002909	-0,03822	8,45E-01	0,05784
b7[1]	-0,02557	0,02023	0,002779	-0,07475	-0,02323	0,003989
b8[1]	-0,8775	0,08503	0,01228	-1,015	-0,8855	-0,6988
b1[2]	0,5267	0,02618	0,003912	0,479	0,5311	0,5612
b2[2]	0,3795	0,01732	0,002565	0,3456	0,3817	0,4089
b3[2]	-1,028	0,0147	0,002109	-1,051	-1,032	-0,9999
b4_1[2]	0,02426	0,07201	0,0107	-0,09264	0,04424	0,1283
b4_2[2]	0,6229	0,0595	0,008888	0,5324	0,613	0,7132
b4_3[2]	-1,329	0,2244	0,02558	-1,825	-1,32	-0,8784
b4_4[2]	-1,766	0,2069	0,02584	-2,241	-1,768	-1,43
b5[2]	-0,8042	0,02204	0,003236	-0,8425	-0,803	-0,769
b6[2]	0,1589	0,005748	0,844	0,1485	0,16	0,1685
b7[2]	0,7912	0,002474	0,361	0,7861	0,7922	0,7945
b8[2]	0,5496	0,00953	0,001406	0,5362	0,5468	0,5681
b1[3]	0,5389	0,01158	0,001692	0,5177	0,538	0,5625
b2[3]	0,2462	0,01038	0,001531	0,2274	0,2447	0,2639
b3[3]	0,1697	0,01448	0,002136	0,1403	0,1742	0,1933
b4_1[3]	0,01034	0,01926	0,002817	-0,01658	0,004454	0,04662
b4_2[3]	0,2694	0,02044	0,003007	0,23	0,2701	0,301
b4_3[3]	-1,024	0,1288	0,01812	-1,266	-1,021	-0,7419
b4_4[3]	0,1204	0,3459	0,05139	-0,5799	0,2927	0,5987
b5[3]	-0,2245	0,01666	0,00247	-0,2509	-0,2278	-0,1948

Lanjutan 1 Tabel 4.5.

Node	Mean	Sd	MC Error	2.5%	Median	97.5%
b6[3]	0,5691	0,002693	0,401	0,5652	0,5688	0,5729
b7[3]	0,3746	0,003618	0,539	0,3694	0,3743	0,3811
b8[3]	0,01988	0,004412	0,632	0,01068	0,02018	0,02699
b1[4]	0,09722	0,03892	0,005682	0,02653	0,09561	0,1833
b2[4]	0,346	0,05114	0,00753	0,2371	0,3584	0,4167
b3[4]	0,05621	0,03718	0,005375	0,001356	0,04833	0,1394
b4_1[4]	1,294	0,1004	0,01461	1,136	1,272	1,492
b4_2[4]	1,262	0,1105	0,01595	1,089	1,232	1,515
b4_3[4]	1,375	0,1336	0,01912	1,149	1,383	1,617
b4_4[4]	1,302	0,1066	0,01574	1,103	1,298	1,514
b5[4]	0,1496	0,08143	0,01203	0,0139	0,1631	0,2691
b6[4]	0,07035	0,002363	0,350	0,06517	0,07105	0,07355
b7[4]	0,02315	0,01196	0,001731	0,001777	0,02325	0,04526
b8[4]	0,0706	0,03307	0,004793	0,01068	0,06714	0,1427
b1[5]	0,1418	0,01107	0,001421	0,1194	0,1418	0,1627
b2[5]	0,2348	0,01508	0,001908	0,203	0,237	0,2607
b3[5]	0,1695	0,009682	0,00108	0,1528	0,1687	0,1882
b4_1[5]	0,603	0,0316	0,004293	0,5353	0,602	0,6664
b4_2[5]	0,6197	0,03245	0,004465	0,5482	0,6193	0,691
b4_3[5]	0,67	0,0315	0,004173	0,6084	0,6691	0,7387
b4_4[5]	0,69	0,02999	0,003995	0,6328	0,69	0,7594
b5[5]	0,2619	0,0141	0,001771	0,2347	0,2604	0,2922
b6[5]	0,01332	5,71E-4	0,0393	0,01206	0,01334	0,01443
b7[5]	0,07189	0,003826	0,508	0,06466	0,07165	0,07988
b8[5]	0,2039	0,0107	0,001354	0,1821	0,2045	0,2215
b1[6]	0,1225	0,02813	0,003531	0,05828	0,124	0,1707
b2[6]	0,7286	0,09177	0,01305	0,5384	0,7199	0,912
b3[6]	0,2573	0,02996	0,003523	0,2015	0,2576	0,311
b4_1[6]	-0,04943	0,08181	0,01005	-0,2002	-0,04617	0,1035
b4_2[6]	0,153	0,06349	0,007874	0,03974	0,1538	0,2768
b4_3[6]	0,09768	0,06401	0,007685	-0,02007	0,09498	0,2263
b4_4[6]	0,05246	0,05544	0,006933	-0,04913	0,04924	0,1588
b5[6]	0,1959	0,03887	0,004955	0,1236	0,1953	0,2728
b6[6]	0,004926	0,002667	0,271	-8,16E-01	0,004963	0,009834
b7[6]	0,05403	0,0124	0,001632	0,02615	0,05466	0,0779
b8[6]	0,5175	0,062	0,008505	0,3849	0,5171	0,6355

Lanjutan 2 Tabel 4.5.

Node	Mean	Sd	MC Error	2.5%	Median	97.5%
ptau[1]	3,66	4,141	0,443	0,1324	2,147	15,8
ptau[2]	0,1962	0,406	0,04869	0,001191	0,04954	1,423
ptau[3]	0,01153	0,006376	0,0007346	0,003676	0,01007	0,02717
ptau[4]	0,2724	0,04062	0,002609	0,1983	0,2693	0,3602
ptau[5]	0,1394	0,003863	0,161	0,1317	0,1394	0,1468
ptau[6]	0,3749	0,05469	0,002521	0,2781	0,3719	0,4909
sigma	0,3562	0,04285	0,00135	0,2837	0,3522	0,4537
dtau	8,216	1,915	0,06194	4,863	8,061	12,45

Sumber : Pengolahan WinBUGS

Seperti yang telah dijelaskan pada bab 3, bahwa variabel-variabel yang diduga berpengaruh terhadap laju kelulusan anak umur 16-18 tahun meliputi jenis kelamin (X1), status perkawinan (X2), status bekerja (X3), tingkat pendidikan kepala rumah tangga (X4), jenis kelamin kepala rumah tangga (X5), rata-rata pengeluaran rumah tangga per kapita (X6), Jumlah anggota rumah tangga (X7), dan lokasi tempat tinggal (kota atau desa) (X8). Hasil estimasi (*posterior summaries*) terhadap variabel-variabel yang mempengaruhi laju kelulusan responden anak umur 16-18 tahun secara lengkap dapat dilihat pada Lampiran 6.

Hasil estimasi parameter (*posterior summaries*) model *mixture survival* spasial bedistribusi Normal disajikan pada Tabel 4.5. Variabel-variabel yang dianggap signifikan mempengaruhi laju kelulusan responden anak umur 16-18 tahun jika nilai pada selang 2,5% hingga 97,5% tidak memuat nilai 0. Tabel 4.5 menunjukkan bahwa tidak semua variabel berpengaruh signifikan terhadap laju kelulusan responden anak umur 16-18 tahun dan terdapat komponen *mixture* yang dipengaruhi oleh bebrapa variabel saja. Kolom *node* merupakan kovariat (variabel-variabel) yang diduga berpengaruh terhadap laju kelulusan responden anak umur 16-18 tahun, kolom *mean* menunjukkan besarnya nilai parameter model, sedangkan ketiga kolom berikutnya besarnya nilai estimasi pada selang kepercayaan 95%.

Pada komponen *mixture* pertama dengan besar proporsi ($\Phi[1]$) 0,00869, responden yang masuk yaitu responden dengan angka lama sekolah 0, 1 dan 2. Variabel-variabel yang secara signifikan mempengaruhi model meliputi status

perkawinan responden (X2), status bekerja responden (X3), jenis kelamin kepala rumah tangga (X5), dan lokasi tempat tinggal responden (X8). Ciri khas dari kelompok *mixture* pertama ini yaitu daerah tempat tinggal responden yang di daerah pedesaan.

Untuk komponen *mixture* kedua dengan besar proporsi ($\Phi[2]$) 0,00767, responden yang masuk yaitu responden dengan angka lama sekolah 1, 2, 3, 4, dan 5, serta variabel-variabel yang secara signifikan mempengaruhi model meliputi semua variabel kecuali variabel dummy kategori 1 pada variabel tingkat pendidikan tertinggi kepala keluarga. Ciri khas dari kelompok *mixture* kedua ini yaitu tingkat pendidikan kepala rumah tangga yang didominasi tidak tamat SD dan lokasi rumah pedesaan.

Untuk komponen *mixture* ketiga dengan besar proporsi ($\Phi[3]$) 0,0642, responden yang masuk yaitu responden dengan angka lama sekolah 4, 5, 6, 7, dan 8. Variabel-variabel yang secara signifikan mempengaruhi model meliputi semua variabel kecuali variabel dummy kategori 1 dan kategori 4 pada variabel tingkat pendidikan tertinggi kepala keluarga, dengan ciri khas dari kelompok *mixture* ketiga ini tingkat pendidikan kepala rumah tangga yang secara signifikan mempengaruhi dan mendominasi yaitu kepala rumah tangga dengan tingkat pendidikan tertinggi tamat SMP

Untuk komponen *mixture* keempat dengan besar proporsi ($\Phi[4]$) 0,09174, responden yang masuk yaitu responden dengan angka lama sekolah 6, 7, 8, 9, dan 10, dimana semua variabel secara signifikan mempengaruhi model pada *mixture* keempat. Ciri khas dari kelompok *mixture* keempat ini yaitu tingkat pendidikan tertinggi kepala rumah tangga yang mempunyai ijazah hampir seimbang dengan yang tidak punya ijazah (tidak tamat SD)

Untuk *mixture* kelima dengan besar proporsi ($\Phi[5]$) 0,7984, responden yang masuk yaitu responden dengan angka lama sekolah 8, 9, 10, 11, dan 12, dimana sama hal nya seperti *mixture* keempat dimana semua variabel secara signifikan mempengaruhi model pada *mixture* kelima. Ciri khas dari kelompok *mixture* kelima ini yaitu tempat tinggal responden yang didominasi pada daerah perkotaan.

Sedangkan untuk *mixture* keenam dengan besar proporsi (Phi[6]) 0,02929, responden yang masuk yaitu responden dengan angka lama sekolah 10, 11, dan 12. Variabel-variabel yang secara signifikan mempengaruhi model pada *mixture* keenam meliputi jenis kelamin responden (X1), status perkawinan responden (X2), status bekerja responden (X3), jenis kelamin kepala rumah tangga (X5), jumlah anggota rumah tangga (X7), dan lokasi tempat tinggal responden (X8). Ciri khas dari kelompok *mixture* keenam ini yaitu hampir semua responden dengan status belum kawin dan mendominasi tempat tinggal di daerah perkotaan serta tingkat pendidikan tertinggi kepala rumah tangga yang didominasi tamat SMA.

Dalam menginterpretasikan tingkat resiko (kecenderungan) suatu variabel digunakan nilai *odds ratio* yang merupakan suatu perbandingan individu pada kondisi variabel/variabel prediktor (x) kategori yang diharapkan dengan variabel/variabel prediktor (x) pada kategori pembandingan.

Tabel 4.6. Odds Ratio

Node	Odds Ratio	Node	Odds Ratio	Node	Odds Ratio
<i>Mixture 1</i>		<i>Mixture 2</i>		<i>Mixture 3</i>	
b2[1]	2,68	b1[2]	1,69	b1[3]	1,71
b3[1]	3,77	b2[2]	1,46	b2[3]	1,28
b5[1]	1,51	b3[2]	0,36	b3[3]	1,18
b8[1]	0,42	b4_2[2]	1,86	b4_2[3]	1,31
		b4_3[2]	0,26	b4_3[3]	0,36
		b4_4[2]	0,17	b5[3]	0,80
		b5[2]	0,45	b6[3]	1,77
		b6[2]	1,17	b7[3]	1,45
		b7[2]	2,21	b8[3]	1,02
		b8[2]	1,73		

Sumber : Pengolahan Excel

Lanjutan Tabel 4.6.

Node	Odds Ratio	Node	Odds Ratio	Node	Odds Ratio
<i>Mixture 4</i>		<i>Mixture 5</i>		<i>Mixture 6</i>	
b1[4]	1,10	b1[5]	1,15	b1[6]	1,13
b2[4]	1,41	b2[5]	1,26	b2[6]	2,07
b3[4]	1,06	b3[5]	1,18	b3[6]	1,29
b4_1[4]	3,65	b4_1[5]	1,83	b4_2[6]	1,17
b4_2[4]	3,53	b4_2[5]	1,86	b5[6]	1,22
b4_3[4]	3,96	b4_3[5]	1,95	b7[6]	1,06
b4_4[4]	3,68	b4_4[5]	1,99	b8[6]	1,68
b5[4]	1,16	b5[5]	1,30		
b6[4]	1,07	b6[5]	1,01		
b7[4]	1,02	b7[5]	1,07		
b8[4]	1,07	b8[5]	1,23		

Sumber : Pengolahan Excel

Berdasarkan Tabel 4.5 dan Tabel 4.6, variabel-variabel yang dinyatakan berpengaruh signifikan terhadap laju kelulusan responden anak umur 16-18 tahun dengan selang kepercayaan 95% beserta interperasinya berdasarkan nilai *odds ratio* sebagai berikut,

1. Variabel Jenis Kelamin Responden (X1)

Variabel jenis kelamin responden (X1) secara signifikan mempengaruhi model *mixture survival* spasial pada *mixture* kedua ($b1[2]=0,5627$), ketiga ($b1[3]=0,5389$), keempat ($b1[4]=0,0972$), kelima ($b1[5]=0,1418$) dan keenam ($b1[6]=0,1225$). Pada variabel jenis kelamin yang menjadi acuan untuk tingkat kecenderungan yaitu responden dengan kategori jenis kelamin laki-laki, sedangkan responden kategori jenis kelamin perempuan sebagai pembanding.

Pada *mixture* kedua, tingkat kecenderungan responden dengan jenis kelamin laki-laki untuk memperoleh kesempatan sekolah lebih lama sebesar 1,69 kali daripada responden dengan jenis kelamin perempuan, artinya jika ada responden perempuan dengan angka lama sekolah 7 tahun maka responden laki-laki memiliki angka lama sekolah minimal 12 tahun.

Pada *mixture* ketiga sebesar 1,71 kali, pada *mixture* keempat sebesar 1,10 kali, pada *mixture* kelima sebesar 1,15 kali, sedangkan pada *mixture* keenam, sebesar 1,13 kali.

2. Variabel Status Perkawinan Responden (X2)

Variabel status perkawinan responden (X2) secara signifikan mempengaruhi model *mixture survival* spasial pada semua komponen *mixture*. Dengan nilai parameter dari variabel status perkawinan responden pada *mixture* pertama (b2[1]) 0,9876, *mixture* kedua (b2[2]) 0,3795, *mixture* ketiga (b2[3]) 0,2462, *mixture* keempat (b2[4]) 0,346, *mixture* kelima (b2[5]) 0,2348 dan *mixture* keenam (b2[6]) 0,7286. Pada variabel status perkawinan yang menjadi acuan untuk tingkat kecenderungan yaitu responden dengan kategori status belum kawin, sedangkan responden kategori status sudah kawin sebagai pembanding.

Pada *mixture* pertama, tingkat kecenderungan responden yang berstatus belum kawin untuk memperoleh kesempatan sekolah lebih lama sebesar 2,68 kali daripada responden yang sudah kawin, artinya jika ada responden dengan status kawin putus sekolah pada angka lama sekolah 4 tahun maka responden yang belum kawin akan memiliki angka lama sekolah minimal 11 tahun. Hal tersebut dikarenakan jika seorang anak kawin pada jenjang pendidikan SD, SMP, atau SMA, anak tersebut dikeluarkan dari sekolah/*drop out*, maka responden tersebut sulit mendapatkan kesempatan sekolah lebih lama.

Pada *mixture* kedua sebesar 1,46 kali, pada *mixture* keempat sebesar 1,41 kali, pada *mixture* kelima sebesar 1,27 kali, sedangkan pada *mixture* keenam sebesar 2,07.

3. Variabel Status Bekerja Responden (X3)

Variabel status bekerja responden (X3) secara signifikan mempengaruhi model *mixture survival* spasial pada semua komponen *mixture*. Dengan nilai parameter dari variabel status bekerja responden pada *mixture* pertama (b3[1]) 1,328, *mixture* kedua (b3[2]) -1,028, *mixture* ketiga (b3[3]) 0,1697, *mixture* keempat (b3[4]) 0,056, *mixture* kelima (b3[5]) 0,1695 dan *mixture* keenam (b3[6]) 0,2573. Pada variabel status bekerja yang menjadi acuan untuk tingkat

kecenderungan yaitu responden dengan kategori status belum bekerja, sedangkan responden kategori status sudah bekerja sebagai pembandingan.

Pada *mixture* pertama, tingkat kecenderungan responden yang berstatus belum bekerja untuk memperoleh kesempatan sekolah lebih lama sebesar $\exp 3,77$ kali daripada responden yang bekerja, artinya jika ada responden dengan status bekerja pada usia sekolah dengan angka lama sekolah 3 tahun maka responden yang belum bekerja akan memiliki angka lama sekolah minimal 11 tahun. Hal tersebut dikarenakan jika seorang anak bekerja pada usia sekolah maka anak tersebut akan lebih cenderung untuk memprioritaskan pekerjaan daripada sekolah.

Pada *mixture* kedua sebesar 0,36 kali, pada *mixture* ketiga sebesar 1,18 kali, pada *mixture* keempat sebesar 1,06 kali, pada *mixture* kelima sebesar 1,18 kali, sedangkan pada *mixture* keenam sebesar 1.29 kali daripada responden yang sudah bekerja.

Kecuali pada *mixture* kedua, pada komponen *mixture* lain responden anak umur 16-18 tahun yang berstatus belum bekerja tingkat kecenderungan memiliki kesempatan sekolah lebih lama diatas 9 tahun (lulus tingkat SMP/ sederajat) lebih tinggi daripada responden yang berstatus sudah bekerja.

4. Variabel Tingkat Pendidikan Tertinggi Kepala Rumah Tangga (X4)

Variabel tingkat pendidikan kepala rumah tangga (X4) dari responden berbentuk data ordinal dengan 5 kategori, seperti yang telah dijelaskan pada bab 2, kategori 1 untuk kepala rumah tangga yang tidak tamat SD/ sederajat, kategori 2 untuk kepala rumah tangga yang tamat SD/ sederajat, kategori 3 untuk kepala rumah tangga yang tamat SMP/ sederajat, kategori 4 untuk kepala rumah tangga yang tamat SMA/ sederajat, dan kategori 5 untuk kepala rumah tangga yang tamat PT. Dalam penentuan variabel dummy untuk variabel tingkat pendidikan kepala rumah tangga, dummy 1 (X4_1) saat kategori 1 (kepala rumah tangga dengan tingkat pendidikan terakhir tidak tamat SD/ sederajat) ada dan kategori yang lain tidak ada, dummy 2 (X4_2) saat kategori 2 (kepala rumah tangga dengan tingkat pendidikan terakhir tamat SD/ sederajat) ada dan kategori yang lain tidak ada, dummy 3 (X4_3) saat kategori 3 (kepala rumah tangga dengan tingkat pendidikan terakhir tamat SMP/ sederajat) ada dan kategori yang lain tidak ada, dummy 4

(X4_4) saat kategori 4 (kepala rumah tangga dengan tingkat pendidikan terakhir tamat SMA/ sederajat) ada dan kategori yang lain tidak ada, dan saat dummy 1 hingga dummy 4 tidak ada, maka itu saat kategori 5 ada.

Variabel tingkat pendidikan kepala rumah tangga (X4) dari responden secara signifikan mempengaruhi model *mixture survival* spasial pada beberapa komponen *mixture* (*mixture* kedua, ketiga, keempat, kelima, dan keenam) dan dummy juga yang signifikan pada tiap komponen ada yang semua dummy, ada juga yang sebagian dikarenakan ada dummy yang tidak memiliki responden pada komponen *mixture* tertentu.

Pada *mixture* kedua dummy yang secara signifikan mempengaruhi dummy 2 ($b4_2[2]=0,6229$), dummy 3 ($b4_3[2]=-1,329$), dan dummy 4 ($b4_4[2]=-1,766$). Pada *mixture* ketiga dummy yang secara signifikan mempengaruhi dummy 2 ($b4_2[3]=0,2694$) dan dummy 3 ($b4_3[3]=-1,024$). Pada *mixture* keempat yang secara signifikan mempengaruhi semua dummy dengan dummy 1 ($b4_1[4]=1,294$), dummy 2 ($b4_2[4]=1,262$), dummy 3 ($b4_3[4]=1,375$), dan dummy 4 ($b4_4[4]=1,302$). Begitu juga pada *mixture* kelima dimana dummy 1 ($b4_1[5]=0,603$), dummy 2 ($b4_2[5]=0,6197$), dummy 3 ($b4_3[5]=0,67$), dan dummy 4 ($b4_4[5]=0,69$). Sedangkan pada *mixture* keenam hanya dummy 2 saja yang signifikan dengan nilai parameter $b4_2[6]=0,153$.

Pada *mixture* kedua, kepala rumah tangga dengan status pendidikan tertinggi tamat SD/ sederajat lebih cenderung memiliki anak umur 16-18 tahun (dalam hal ini responden) yang memperoleh kesempatan bersekolah lebih lama sebesar 1,86 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain. Untuk kepala rumah tangga dengan status pendidikan tertinggi tamat SMP/ sederajat lebih cenderung memiliki anak umur 16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 0,26 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain, sedangkan untuk kepala rumah tangga dengan status pendidikan tertinggi tamat SMA/ sederajat lebih cenderung memiliki anak umur 16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 0,17 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain.

Pada *mixture* ketiga, artinya kepala rumah tangga dengan status pendidikan tertinggi tamat SD/ sederajat lebih cenderung memiliki anak umur 16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 1,31 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain, sedangkan untuk kepala rumah tangga dengan status pendidikan tertinggi tamat SMP/ sederajat lebih cenderung memiliki anak umur 16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 0,36 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain.

Pada *mixture* keempat, kepala rumah tangga dengan status pendidikan tertinggi tidak tamat SD/ sederajat lebih cenderung memiliki anak umur 16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 3,65 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain, untuk kepala rumah tangga dengan status pendidikan tertinggi tamat SD/ sederajat lebih cenderung memiliki anak umur 16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 3,53 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain, untuk kepala rumah tangga dengan status pendidikan tertinggi tamat SMP/ sederajat lebih cenderung memiliki anak umur 16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 3,96 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain, sedangkan untuk kepala rumah tangga dengan status pendidikan tertinggi tamat SMA/ sederajat lebih cenderung memiliki anak umur 16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 3,68 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain.

Pada *mixture* kelima, kepala rumah tangga dengan status pendidikan tertinggi tidak tamat SD/ sederajat lebih cenderung memiliki anak umur 16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 1,83 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain, untuk kepala rumah tangga dengan status pendidikan tertinggi tamat SD/ sederajat lebih cenderung memiliki anak umur 16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 1,86 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain, untuk kepala rumah tangga dengan status pendidikan tertinggi tamat SMP/ sederajat lebih cenderung memiliki anak umur

16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 1,95 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain, sedangkan untuk kepala rumah tangga dengan status pendidikan tertinggi tamat SMA/ sederajat lebih cenderung memiliki anak umur 16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 1,95 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain.

Pada *mixture* keenam, variabel pendidikan tertinggi kepala rumah tangga dari responden untuk kepala rumah tangga dengan status pendidikan tertinggi tamat SD/ sederajat lebih cenderung memiliki anak umur 16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 1,17 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain.

5. Variabel Jenis Kelamin Kepala Rumah Tangga (X5)

Variabel jenis kelamin kepala rumah tangga (X5) secara signifikan mempengaruhi model *mixture survival* spasial pada semua komponen *mixture*. Dengan nilai parameter dari variabel jenis kelamin kepala rumah tangga responden pada *mixture* pertama (b5[1]) 0,4103, *mixture* kedua (b5[2]) -0,8042, *mixture* ketiga (b5[3]) -0,2245, *mixture* keempat (b5[4]) 0,1496, *mixture* kelima (b5[5]) 0,2619 dan *mixture* keenam (b5[6]) 0,1959. Kategori yang menjadi acuan sama seperti variabel X1 yaitu kepala rumah tangga berjenis kelamin laki-laki, sedangkan kepala rumah tangga berjenis kelamin perempuan sebagai pembanding.

Pada *mixture* pertama, tingkat kecenderungan kepala rumah tangga dengan jenis kelamin laki-laki memiliki anak umur 16-18 tahun untuk memperoleh kesempatan bersekolah lebih lama sebesar 1,51 kali daripada responden dengan kepala rumah tangga berjenis kelamin perempuan, artinya jika ada responden dengan kepala rumah tangga perempuan memiliki angka lama sekolah 7 tahun, maka responden lainnya dengan kepala rumah tangga laki-laki akan memiliki angka lama sekolah minimal sebesar 8 tahun. Responden dengan kepala rumah tangga perempuan akan lebih cenderung untuk bekerja juga membantu keuangan rumah tangga sehingga untuk keperluan sekolah dikesampingkan.

Pada *mixture* kedua sebesar 0,45 kali, pada *mixture* ketiga sebesar 0,8 kali, pada *mixture* keempat sebesar 1,16 kali, pada *mixture* kelima sebesar 1,3 kali,

sedangkan pada *mixture* keenam sebesar 1,22 kali daripada responden dengan kepala rumah tangga berjenis kelamin perempuan.

6. Variabel Pengeluaran Per Kapita Per Bulan Rumah Tangga (X6)

Variabel pengeluaran per kapita per bulan rumah tangga (X6) secara signifikan mempengaruhi model *mixture survival* spasial pada *mixture* kedua ($b6[2]=0,1589$), *mixture* ketiga ($b6[3]=0,5691$), *mixture* keempat ($b6[4]=0,07035$), dan *mixture* kelima ($b6[5]=0,01332$). Seperti yang telah dijelaskan pada bab 2 bahwa interpretasi untuk variabel kontinu (data rasio dan interval) berbeda dengan interpretasi untuk variabel kategorik.

Pada *mixture* kedua, setiap penambahan 1 satuan pengeluaran per kapita per bulan rumah tangga lebih cenderung seorang responden memperoleh kesempatan bersekolah lebih lama sebesar 1,17 kali atau dalam interpretasi lain, setiap penambahan 1 s atuan pengeluaran per kapita per bulan rumah tangga lebih cenderung membuat keinginan anak umur 16-18 tahun pada rumah tangga tersebut sebesar 1,17 kali untuk tetap bersekolah.

Untuk *mixture* ketiga sebesar 1,77 kali, pada *mixture* keempat sebesar 1,07 kali, sedangkan pada *mixture* kelima sebesar sebesar 1,01 kali

7. Variabel Jumlah Anggota Rumah Tangga (X7)

Variabel jumlah anggota rumah tangga (X7) secara signifikan mempengaruhi model *mixture survival* spasial pada *mixture* kedua ($b7[2]=0,7912$), *mixture* ketiga ($b7[3]=0,3746$), *mixture* keempat ($b7[4]=0,0232$), *mixture* kelima ($b7[5]=0,07189$), dan *mixture* keenam ($b7[6]=0,054$).

Pada *mixture* kedua, setiap penambahan 1 s atuan anggota rumah tangga lebih cenderung membuat keinginan anak umur 16-18 tahun pada rumah tangga tersebut sebesar 2.21 kali untuk mendapatkan kesempatan bersekolah lebih sedikit atau tidak melanjutkan pendidikan untuk jenjang pendidikan yang lebih tinggi dari SMP/ sederajat.

Pada *mixture* ketiga sebesar 1.45 kali, pada *mixture* keempat sebesar 1.02 kali, pada *mixture* kelima sebesar 1.07 kali, sedangkan pada *mixture* keenam sebesar 1.06 kali.

8. Variabel Lokasi Tempat Tinggal Responden (X8)

Variabel lokasi tempat tinggal responden (X8) yaitu perkotaan atau perdesaan secara signifikan mempengaruhi model *mixture survival* spasial pada semua komponen *mixture*. Dengan nilai parameter dari variabel lokasi tempat tinggal responden pada *mixture* pertama (b8[1]) -0,8775, *mixture* kedua (b8[2]) 0,5496, *mixture* ketiga (b8[3]) 0,02, *mixture* keempat (b8[4]) 0,07, *mixture* kelima (b8[5]) 0,2039 dan *mixture* keenam (b8[6]) 0,5175. Kategori acuan yaitu responden yang berlokasi tempat tinggal di perkotaan, sedangkan responden yang berlokasi tempat tinggal di perdesaan menjadi pembandingan.

Pada *mixture* pertama, tingkat kecenderungan responden dengan lokasi tempat tinggal di perkotaan untuk memperoleh kesempatan sekolah lebih lama sebesar 0,42 kali daripada responden dengan lokasi tempat tinggal di perdesaan. Pada *mixture* kedua sebesar 1,73 kali, pada *mixture* ketiga sebesar 1,02 kali, pada *mixture* keempat sebesar 1,07 kali, pada *mixture* kelima sebesar 1,23 kali, sedangkan pada *mixture* keenam sebesar 1,68 kali daripada responden dengan lokasi tempat tinggal di perdesaan.

Selain variabel-variabel yang telah disebutkan di atas, dapat diketahui bahwa parameter *frailty* CAR ($\sigma(\lambda) = 0.3562$) juga signifikan berpengaruh terhadap laju kelulusan responden anak umur 16-18 tahun, dapat dilihat. Hal ini berarti bahwa dalam kasus angka lama sekolah ini memang terdapat efek spasial dalam model *mixture survival* standar sehingga mengakibatkan munculnya efek spasial.

Tabel 4.7. Efek Spasial dari Setiap Kabupaten di Jawa Timur

Node	Mean	Sd	Mc Error	2.5%	Median	97.5%
W[1]	-0.2717	0.04546	0.006049	-0.3782	-0.2672	-0.1914
W[2]	-0.3483	0.03787	0.005593	-0.4128	-0.3565	-0.285
W[3]	0.00197	0.03312	0.002512	-0.05956	0.001607	0.06699
W[4]	-0.07598	0.02272	0.002929	-0.1185	-0.07568	-0.02693
W[5]	-0.02175	0.03081	0.004051	-0.07908	-0.0228	0.04001
W[6]	-0.009028	0.02444	0.002798	-0.05653	-0.008721	0.0343
W[7]	-0.02994	0.02822	0.003328	-0.07783	-0.03216	0.02657
W[8]	0.05817	0.03094	0.003085	9,87E-01	0.05759	0.1213
W[9]	0.0655	0.02438	0.003243	0.02274	0.065	0.1122
W[10]	-0.01428	0.03373	0.004913	-0.06975	-0.02132	0.06326
W[11]	0.04552	0.03446	0.003978	-0.0161	0.04428	0.1211
W[12]	0.0855	0.03918	0.004739	0.01705	0.08556	0.1652
W[13]	0.00452	0.02288	0.002952	-0.04237	0.004623	0.05193
W[14]	-0.05524	0.03654	0.005401	-0.1101	-0.06866	0.009866
W[15]	0.08199	0.0243	0.001635	0.0331	0.08204	0.1291
W[16]	-0.009999	0.02181	0.002487	-0.05154	-0.01001	0.03211
W[17]	0.0905	0.02432	0.0021	0.04211	0.08982	0.1366
W[18]	-0.0434	0.02463	0.003543	-0.08477	-0.04538	-0.003292
W[19]	0.01097	0.02669	0.001846	-0.04092	0.007773	0.07024
W[20]	0.04101	0.0381	0.002447	-0.03863	0.04283	0.1117
W[21]	0.03902	0.02956	0.00267	-0.01865	0.03782	0.096
W[22]	-0.2173	0.0168	0.002384	-0.249	-0.2142	-0.1918
W[23]	-0.01373	0.0255	0.002468	-0.06406	-0.01291	0.03949
W[24]	-0.1714	0.02835	0.004137	-0.2106	-0.1778	-0.1147
W[25]	0.04598	0.02255	0.001923	0.001085	0.04582	0.09086
W[26]	-0.08357	0.03369	0.00477	-0.1602	-0.08514	-0.023
W[27]	-0.1708	0.02792	0.003044	-0.2244	-0.1722	-0.1075
W[28]	-0.06221	0.02883	0.002299	-0.1202	-0.06231	-0.006239
W[29]	0.03104	0.03067	0.001958	-0.02994	0.03154	0.09379
W[30]	0.137	0.03376	0.00326	0.07085	0.1384	0.2001
W[31]	0.1149	0.03634	0.001972	0.04691	0.1154	0.1807
W[32]	0.1578	0.02699	0.003622	0.1054	0.1563	0.2083
W[33]	0.1685	0.0288	0.001633	0.1138	0.1675	0.2249
W[34]	0.1268	0.02978	0.00356	0.07493	0.1266	0.1816
W[35]	0.1157	0.03315	0.003548	0.05803	0.1146	0.1899
W[36]	0.12	0.0342	0.002553	0.04685	0.1217	0.1847
W[37]	0.07523	0.02374	0.00184	0.03056	0.07608	0.1176
W[38]	-0.01891	0.03777	0.003184	-0.09832	-0.02053	0.05435

Sumber : Perhitungan WinBUGS

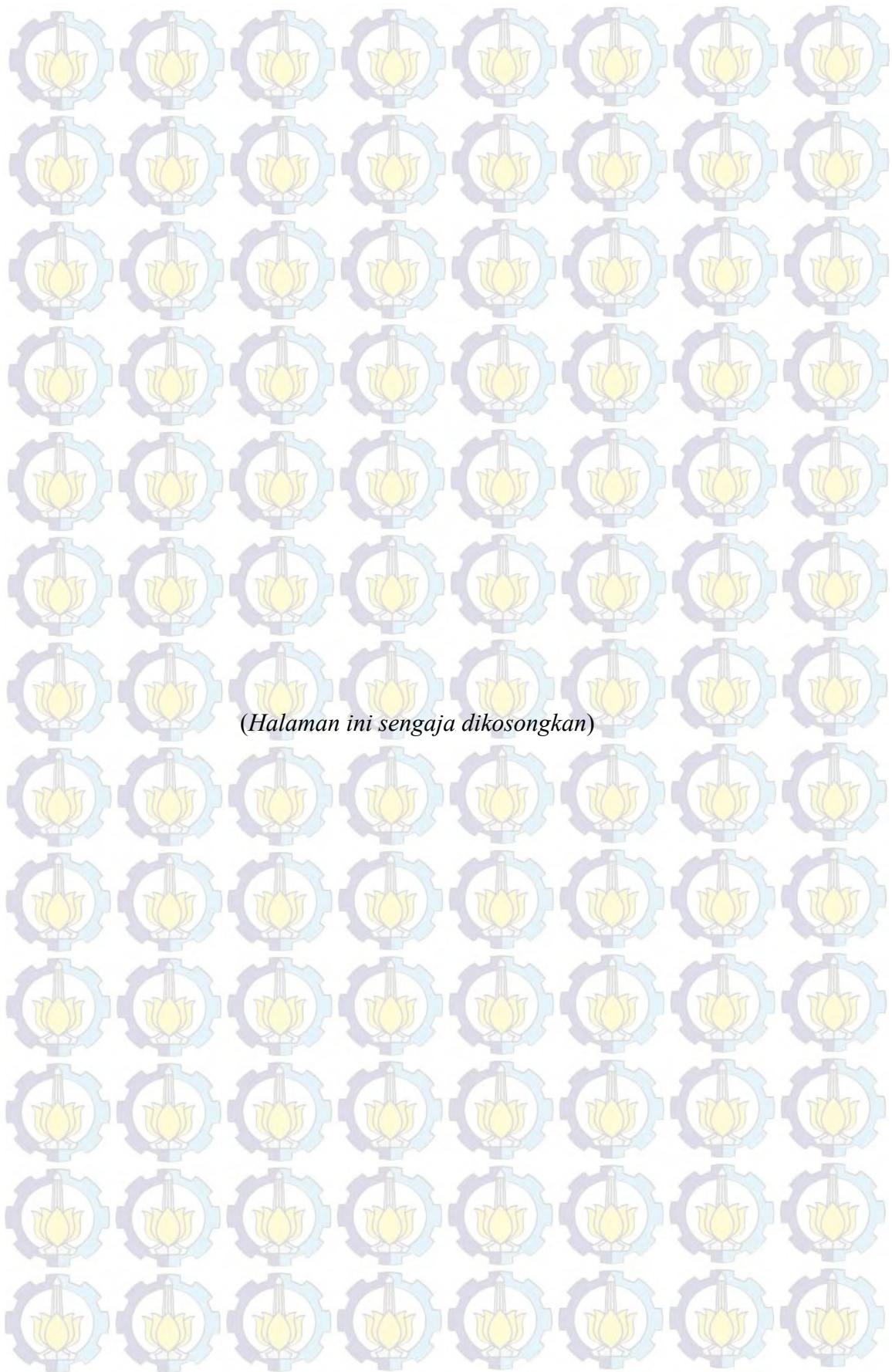
Pada Tabel 4.7. dapat dilihat bahwa baris yang berwarna kuning adalah daerah-daerah dengan efek spasial yang tidak signifikan, daerah-daerah tersebut adalah Trenggalek, Blitar, Kediri, Malang, Banyuwangi, Bondowoso, Probolinggo, Pasuruan, Mojokerto, Madiun, Magetan, Ngawi, Tuban, Sumenep, dan Kota Batu. sedangkan yang berwarna putih adalah daerah-daerah efek spasial yang signifikan, daerah-daerah tersebut yaitu Pacitan, Ponorogo, Tulungagung, Lumajang, Jember, Situbondo, Sidoarjo, Jombang, Nganjuk, Bojonegoro, Lamongan, Gresik, Bangkalan, Sampang, Pamekasan, Kota Kediri, Kota Blitar, Kota Malang, Kota Probolinggo, Kota Pasuruan, Kota Mojokerto, Kota Madiun, dan Kota Surabaya. Signifikan atau tidaknya efek spasial dapat dilihat dari nilai pada selang 2.5% hingga 97.5% tidak memuat nilai 0.

Daerah-daerah dengan efek spasial yang secara signifikan mempengaruhi angka lama sekolah anak umur 16-18 tahun artinya bahwa pengaruh spasial atau pengaruh kedaerahan, dimana dalam penelitian ini pengaruh spasial dianggap random (menjadi variabel random W_j^*), memiliki efek yang berbeda di setiap daerah dan setiap daerah mungkin saja memiliki kebijakan yang berbeda atau keadaan yang berbeda yang diciptakan oleh masyarakat antara satu daerah dengan daerah lain sehingga mempengaruhi angka lama sekolah anak umur 16-18 tahun di daerah tersebut yang berimplikasi bahwa rata-rata lama sekolah berdasarkan angka lama sekolah anak umur 16-18 tahun di daerah tersebut berbeda juga.

Berdasarkan Tabel 4.5 dan Tabel 4.7, maka model mixture survival spasial yang terbentuk yaitu

$$\begin{aligned}
h(t, \mathbf{x}) = & 0,0087(h_1(t) \exp(0,98x_2 + 1,33x_3 + 0,41x_5 + 0,88x_8 + W_j^*)) \\
& + 0,008(h_2(t) \exp(0,53x_1 + 0,38x_2 - 1,03x_3 + 0,62x_{4_2} \\
& - 1,33x_{4_3} - 1,77x_{4_4} - 0,8x_5 + 0,16x_6 + 0,8x_7 + 0,55x_8 + W_j^*)) \\
& + 0,06(h_3(t) \exp(0,54x_1 + 0,25x_2 + 0,17x_3 + 0,27x_{4_2} - 1,02x_{4_3} \\
& - 0,22x_5 + 0,57x_6 + 0,37x_7 + 0,02x_8 + W_j^*)) \\
& + 0,092(h_4(t) \exp(0,1x_1 + 0,34x_2 + 0,06x_3 + 1,29x_{4_1} + 1,26x_{4_2} \\
& + 1,38x_{4_3} + 1,3x_{4_4} + 0,15x_5 + 0,07x_6 + 0,02x_7 + 0,07x_8 + W_j^*)) \\
& + 0,798(h_5(t) \exp(0,14x_1 + 0,23x_2 + 0,17x_3 + 0,6x_{4_1} + 0,62x_{4_2} \\
& + 0,67x_{4_3} + 0,69x_{4_4} + 0,26x_5 + 0,013x_6 + 0,072x_7 + 0,2x_8 + W_j^*)) \\
& + 0,029(h_6(t) \exp(0,122x_1 + 0,73x_2 + 0,26x_3 + 0,15x_{4_2} + 0,2x_5 \\
& + 0,05x_7 + 0,52x_8 + W_j^*))
\end{aligned}$$

dimana $h_1(t)$ hingga $h_6(t)$ adalah fungsi baseline hazard dari setiap komponen mixture dan W_j^* adalah efek spasial dari daerah ke- j . Nilai $h(t, \mathbf{x})$ adalah nilai besarnya laju atau kesempatan untuk bersekolah lebih lama dengan dipengaruhi oleh besar komponen, nilai baseline hazard, variabel-variabel yang secara signifikan mempengaruhi pada komponen tertentu dan daerah yang memiliki efek spasial yang secara signifikan mempengaruhi.



BAB 4

HASIL DAN PEMBAHASAN

Pada bab ini akan dijabarkan tentang hal mendasar dalam penelitian beserta tahapan-tahapan dalam menjalankannya, diantaranya sumber data, variabel penelitian, metode dan tahapan penelitian serta jadwal penelitian.

4.1. Model *Survival* Dengan Penambahan Konsep *Mixture* dan Pertimbangan Adanya Efek Spasial Menggunakan Pendekatan Bayesian.

Pada bagian ini akan menjawab tujuan pertama penelitian yaitu mengenai penjabaran model *survival* dengan penambahan konsep *mixture* dan pertimbangan adanya efek spasial menggunakan pendekatan bayesian, yang meliputi hubungan model *survival* dan model *hazard* serta didapatnya mode *hazard* proporsional, penambahan efek random (*frailty*) dalam model *hazard* proporsional, penambahan efek random spasial (*spatial frailty*) dalam model *hazard* proporsional, penentuan distribusi *prior* dan *joint* distribusi *posterior* dari model *mixture survival* spasial berdasarkan distribusi data, serta estimasi parameter model *mixture survival* spasial menggunakan MCMC dan *gibbs sampling*.

4.1.1. Hubungan Antara Fungsi *Survival* Dengan Fungsi *Hazard*

Dalam analisis *survival*, terdapat dua fungsi yang digunakan yaitu fungsi *survival* dan fungsi *hazard*. Pada bab 2 telah dituliskan bahwa Le (1997) mendefinisikan fungsi *survival* $S(t)$ sebagai probabilitas seorang individu bertahan lebih lama dari waktu t . Fungsi *survival* dapat dilihat pada persamaan (2.3).

Lee (1992) mendefinisikan fungsi *hazard* sebagai peluang kegagalan individu untuk bertahan selama interval waktu yang sangat pendek dengan asumsi bahwa individu tersebut telah bertahan pada awal interval atau limit peluang individu gagal bertahan dalam sebuah interval waktu yang sangat pendek, yaitu dari t sampai Δt jika individu tersebut telah bertahan sampai waktu t seperti yang dituliskan pada persamaan (2.4). Semakin besar nilai *hazard* mengindikasikan

bahwa risiko kegagalan yang dialami individu dalam penelitian semakin tinggi sehingga kemampuan bertahanannya semakin kecil.

Hubungan antara fungsi *survival* dan fungsi *hazard* dapat diperoleh dengan menggunakan teori probabilitas bersyarat, yaitu $P(A|B) = \frac{P(AB)}{P(B)}$, dengan pembilang dalam limit persamaan (2.4) dapat ditentukan hubungannya sebagai berikut:

$$\frac{P(t \leq T < t + \Delta t)}{P(T \geq t)} = \frac{F(t + \Delta t) - F(t)}{S(t)} \quad (4.1)$$

dengan $F(t)$ merupakan fungsi distribusi dari T maka persamaan (2.4) dapat dituliskan:

$$h(t) = \lim_{\Delta t \rightarrow 0} \left\{ \frac{F(t + \Delta t) - F(t)}{\Delta t} \right\} \frac{1}{S(t)} \quad (4.2)$$

dengan mengambil $F'(t) = \lim_{\Delta t \rightarrow 0} \left\{ \frac{F(t + \Delta t) - F(t)}{\Delta t} \right\}$ yang merupakan derivatif pertama dari $F(t)$ maka diperoleh hubungan antara fungsi *survival* dan fungsi *hazard* sebagai berikut:

$$h(t) = \frac{f(t)}{S(t)} \quad (4.3)$$

Diketahui bahwa $F(t) = 1 - S(t)$ dapat dituliskan sebagai $\int f(t)dt = 1 - S(t)$. Jika keduanya diturunkan terhadap t maka diperoleh $f(t) = \frac{d(1 - S(t))}{dt}$ dan nilai $h(t)$ menjadi:

$$h(t) = \frac{\left(\frac{d(1 - S(t))}{dt} \right)}{S(t)} = \frac{\left(-\frac{d}{dt} S(t) \right)}{S(t)} \quad (4.4)$$

$$-h(t)dt = \frac{d(S(t))}{S(t)}$$

dengan mengintegralkan persamaan (4.4), diperoleh

$$-\int_0^t h(t)dt = \int_0^t \frac{1}{S(t)} d(S(t))$$

$$-\int_0^t h(t)dt = \ln S(t) \Big|_0^t = \ln S(t) - \ln S(0) = \ln S(t)$$

$$S(t) = \exp \left[-\int_0^t h(t)dt \right] \quad (4.5)$$

dengan fungsi kumulatif *hazard* adalah sebagai berikut:

$$H(t) = \int_0^t h(t)dt \quad (4.6)$$

sehingga didapatkan hubungan antara fungsi kumulatif *hazard* atau $H(t)$ dengan fungsi *survival* atau $S(t)$ sebagai berikut,

$$H(t) = -\ln S(t) \quad (4.7)$$

Jika resiko gagal (*failure*) pada waktu tertentu bergantung pada nilai x_1, x_2, \dots, x_p dari p variabel prediktor X_1, X_2, \dots, X_p maka nilai variabel tersebut diasumsikan telah tercatat sebagai *time origin*. Kumpulan nilai variabel prediktor dalam model *hazard proporsional* diwakili oleh vektor \mathbf{x} dengan $\mathbf{x} = (x_1, x_2, \dots, x_p)^T$. Misalkan $h_0(t)$ sebagai fungsi *hazard* untuk setiap obyek dengan nilai dari semua variabel prediktor penyusun vektor \mathbf{x} adalah nol maka fungsi $h_0(t)$ dikatakan sebagai fungsi *baseline hazard* (Collet, 2003). Model *hazard proporsional* dinyatakan seperti pada persamaan (2.8), namun untuk mengingatkan dituliskan kembali ke dalam nomor persamaan baru,

$$h(t, \mathbf{x}) = h_0(t) \exp(\mathbf{x}^T \boldsymbol{\beta}) \quad (4.8)$$

4.1.2. Penambahan efek random spasial (*spatial frailty term*) pada model *hazard proporsional*

Pada bab 2 telah dijelaskan bagaimana penambahan efek spasial dan dengan disertakan efek random pada model proporsional *hazard* yang tertulis pada persamaan (2.14), sedangkan untuk model *survival* dengan penambahan efek spasial pada persamaan disertakan efek random (2.16) dan telah dijelaskan juga pada bab 2 bahwa W_j pada persamaan (2.14) dan (2.16) merupakan bentuk *stratum-specific frailty* yang dibentuk untuk menyatakan perbedaan antar strata/kelompok. Secara umum W_j diasumsikan identik, independen dan

berdistribusi normal dengan mean bernilai nol dan varian σ^2 , akan tetapi menurut Hougaard (2000) dalam Banerjee *et al.*, (2003), distribusi non-normal juga sering digunakan untuk memodelkan *frailties*. Artinya tidak harus distribusi normal untuk asumsi W_j .

Model pada pendekatan *survival* spasial dibentuk dari data *survival* yang tersusun secara spasial, artinya *frailties* W_j dari kelompok/strata yang saling berdekatan menggambarkan kemungkinan bahwa kelompok/strata tersebut memiliki nilai/karakteristik yang serupa. Pendekatan lattice digunakan untuk menyusun susunan spasial dari kelompok-kelompok tersebut. Sehingga W_j terganti oleh W_j^* , hal tersebut dikarenakan W^* didefinisikan oleh daerah-daerah dengan indeks diskret dimana daerah-daerah tersebut merupakan partisi/bagian dari daerah D . Sehingga model *hazard* proporsional yang ditambahkan efek spasial menjadi sebagai berikut :

$$h(t_{ij}; \mathbf{x}_{ij}) = h_0(t_{ij}) \exp(\beta_1 x_{1ij} + \beta_2 x_{2ij} + \dots + \beta_p x_{pij} + W_j^*) \quad (4.9)$$

atau dalam bentuk lainnya

model *hazard* proporsional unit i daerah ke-1

$$h(t_{i1}; \mathbf{x}_{i1}) = h_0(t_{i1}) \exp(\beta_1 x_{1i1} + \beta_2 x_{2i1} + \dots + \beta_p x_{pi1} + W_1^*)$$

model *hazard* proporsional unit i daerah ke-2

$$h(t_{i2}; \mathbf{x}_{i2}) = h_0(t_{i2}) \exp(\beta_1 x_{1i2} + \beta_2 x_{2i2} + \dots + \beta_p x_{pi2} + W_2^*)$$

⋮

model *hazard* proporsional unit i daerah ke- j

$$h(t_{ij}; \mathbf{x}_{ij}) = h_0(t_{ij}) \exp(\beta_1 x_{1ij} + \beta_2 x_{2ij} + \dots + \beta_p x_{pij} + W_j^*)$$

⋮

model *hazard* proporsional unit i daerah ke- m

$$h(t_{im}; \mathbf{x}_{im}) = h_0(t_{im}) \exp(\beta_1 x_{1im} + \beta_2 x_{2im} + \dots + \beta_p x_{pim} + W_m^*) \quad (4.10)$$

dengan : x_1, x_2, \dots, x_p adalah variabel prediktor ke-1 hingga ke- p

i menyatakan unit pengamatan, dimana unit pengamatan setiap lokasi berbeda jumlahnya.

j menyatakan daerah lokasi hingga lokasi ke- m

dikarenakan dalam penelitian ini terdapat 8 variabel prediktor dan 38 daerah kabupaten/kota, maka model umum *hazard* proporsional yang ditambahkan efek spasial menjadi sebagai berikut :

$$\begin{aligned} h(t_{i1}; \mathbf{x}_{i1}) &= h_0(t_{i1}) \exp(\beta_1 x_{1i1} + \beta_2 x_{2i1} + \dots + \beta_8 x_{8i1} + W_1^*) \\ &\vdots \\ h(t_{i38}; \mathbf{x}_{i38}) &= h_0(t_{i38}) \exp(\beta_1 x_{1i38} + \beta_2 x_{2i38} + \dots + \beta_8 x_{8i38} + W_{38}^*) \end{aligned} \quad (4.11)$$

4.1.3. Distribusi *prior* dan *joint* distribusi *posterior* dari model *mixture survival* spasial

Distribusi CAR (*Conditionally Autoregressive*) digunakan sebagai distribusi prior untuk parameter v yang mewakili efek random (*frailty*) yang saling berkorelasi secara spasial, dapat dinyatakan sebagai berikut,

$$W^* | \lambda \sim \text{CAR}(\lambda) \quad (4.12)$$

dimana $\lambda = 1/v$ sedangkan bentuk umum dari prior CAR yang memiliki join distribusi proporsional sebagai berikut,

$$\lambda^{1/2} \exp \left[-\frac{\lambda}{2} \sum_{j \text{ adj } j'} (W_j^* - W_{j'}^*)^2 \right] \propto \lambda^{1/2} \exp \left[-\frac{\lambda}{2} \sum_{i=1}^I m_j W_j^* (W_j^* - \bar{W}_j^*) \right] \quad (4.13)$$

dimana : j dan j' adalah wilayah yang saling bertetangga

$$\bar{W}_j^* = \frac{\sum_{j'=1}^I W_{j'}^*}{m_j} \text{ yang merupakan rata-rata dari } W_{j'}^* \text{ yang bertetanggan dengan } W_j^*$$

m_j adalah jumlah tetangga yang dimiliki oleh daerah j

λ juga disebut sebagai parameter penghalus yang secara implisit juga menentukan variabilitas dari efek random (*frailty*). Nilai λ akan berbeda antara daerah satu dengan daerah yang lain, hal ini bergantung pada jumlah tetangga (kedekatan dengan daerah lain) yang dimiliki oleh daerah tersebut (m_j).

Pada penelitian ini, distribusi dari angka lama sekolah anak umur 16-18 tahun mengikuti sebenarnya berdistribusi nonnegatif distribusi karena berbentuk skew kiri, akan tetapi dalam pendekatan untuk setiap komponen *mixture* dari angka lama sekolah anak umur 16-18 tahun dilakukan dengan menggunakan

distribusi Normal 2-parameter (μ, σ) (penjelasan lebih lengkap pada subsub bab 4.2.5) sehingga fungsi kepadatan peluang (PDF) dari normal sebagai berikut,

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) \quad (4.14)$$

sedangkan fungsi distribusi kumulatifnya (CDF) sebagai berikut,

$$\begin{aligned} F(t) = P(T \leq t) &= \int_{-\infty}^t \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) dt \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^t \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) d(t) \end{aligned}$$

misalkan

$$\begin{aligned} \frac{t-\mu}{\sigma} &= x \\ t &= x\sigma + \mu \\ dt &= \sigma dx \end{aligned}$$

maka

$$\begin{aligned} F(t) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^t \exp\left(-\frac{1}{2}(x)^2\right) \sigma dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 \exp\left(-\frac{1}{2}x^2\right) d(x) + \frac{1}{\sqrt{2\pi}} \int_0^t \exp\left(-\frac{1}{2}x^2\right) d(x) \\ &= \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{t-\mu}{\sqrt{2}\sigma}\right) \end{aligned} \quad (4.15)$$

Nilai t selalu bernilai positif dikarenakan waktu survival angka lama dari individu untuk bertahan dari suatu kondisi maka selalu bernilai positif. Pada hasil CDF pada 4.15 terdapat hasil erf dimana erf merupakan *error function* dari CDF

distribusi Normal dikarenakan $\frac{1}{\sqrt{2\pi}} \int_0^t \exp\left(-\frac{1}{2}x^2\right) d(x)$ memiliki hasil yang tidak *close form* sehingga terbentuklah *error function*.

Berdasarkan persamaan CDF pada (4.15) dan fungsi *survival* pada persamaan (2.3), maka dapat ditentukan fungsi *survival* dari distribusi Normal 2-parameter sebagai berikut,

$$S(t) = 1 - F(t) = 1 - \left(\frac{1}{2} + \frac{1}{2} \operatorname{erf} \left(\frac{t - \mu}{\sqrt{2\sigma^2}} \right) \right) = \frac{1}{2} - \frac{1}{2} \operatorname{erf} \left(\frac{t - \mu}{\sqrt{2\sigma^2}} \right) \quad (4.16)$$

Untuk menentukan fungsi *hazard* dari distribusi Normal 2-parameter maka diperlukan PDF pada persamaan (4.14) dan fungsi *survival* pada persamaan (4.16), selanjutnya kedua persamaan tersebut dioperasikan melalui persamaan (4.3) sebagai berikut,

$$h(t) = \frac{f(t)}{S(t)} = \frac{\frac{1}{\sigma\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{t - \mu}{\sigma} \right)^2 \right)}{\frac{1}{2} - \frac{1}{2} \operatorname{erf} \left(\frac{t - \mu}{\sqrt{2\sigma^2}} \right)} = \frac{\sqrt{2} \exp \left(-\frac{1}{2} \left(\frac{t - \mu}{\sigma} \right)^2 \right)}{\sqrt{\pi\sigma^2} \left(1 - \operatorname{erf} \left(\frac{t - \mu}{\sqrt{2\sigma^2}} \right) \right)} \quad (4.17)$$

Berdasarkan fungsi *hazard* pada persamaan (4.17) dan (2.9) maka dapat diperoleh persamaan berikut,

$$h(t, \mathbf{x}) = h_0(t) \exp(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p) = \frac{\sqrt{2} \exp \left(-\frac{1}{2} \left(\frac{t - \mu}{\sigma} \right)^2 \right)}{\sqrt{\pi\sigma^2} \left(1 - \operatorname{erf} \left(\frac{t - \mu}{\sqrt{2\sigma^2}} \right) \right)} \quad (4.18)$$

Pada persamaan (4.18), dikarenakan masih terdapat *error function*, cukup susah untuk menentukan mana yang merupakan bagian dari $h_0(t)$ tetapi untuk $\exp(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)$ dituliskan sebagai berikut :

$$\exp(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p) = \mu \quad (4.19)$$

Sebelum melakukan estimasi terhadap parameter μ dan σ (pada syntax simbol μ disebut psi dan simbol σ disebut tau), maka terlebih dahulu menentukan distribusi priornya sebagai berikut,

$$\sigma \sim \text{Gamma}(p, q)$$

$$\beta_i \sim \text{Normal}(m, s)$$

dimana penentuan dimana penentuan distribusi prior tersebut dilakukan berdasarkan gabungan antara *prior conjugate* dan *prior informatif* yaitu *prior* yang ditentukan berdasarkan pola *likelihood* data dan berdasarkan informasi *prior* dari penelitian sebelumnya. Pada analisis *survival* menyertakan data tersensor dalam menentukan join distribusi posteriornya, γ_{ij} menyatakan status tersensor dari responden angka lama sekolah yaitu nilai 0 untuk responden tersensor dan

nilai 1 untuk responden tidak tersensor. Selanjutnya ditentukan t sebagai waktu *survival* responden (waktu yang diperlukan hingga terjadinya lulus) dan x merupakan vektor *covariate* (variabel prediktor) sehingga diperoleh join distribusi posteriornya sebagai berikut,

$$P(\beta, W, \sigma, \lambda | t, x, \gamma) \propto L(\beta, W, \sigma; t, x, \gamma) p(W | \lambda) p(\beta) p(\sigma) p(\lambda) \quad (4.20)$$

Pada ruas kanan terdapat lima bentuk, yaitu bentuk pertama merupakan likelihood untuk *hazard* distribusi Normal 2-parameter, bentuk kedua menyatakan join distribusi dari efek random (*frailty*), sedangkan tiga bentuk sisanya merupakan distribusi prior dari masing-masing parameter.

Fungsi likelihood untuk *hazard* distribusi Normal 2-parameter dapat dijabarkan sebagai berikut (Collet, 2003),

$$L(\beta, W, \sigma; t, x, \gamma) = \prod_{i=1}^I \prod_{j=1}^{n_i} f_{ij}(t_{ij})^{\gamma_{ij}} S_{ij}(t_{ij})^{1-\gamma_{ij}} \quad (4.21)$$

Dengan menggunakan persamaan (4.4) sehingga dapat ditulis $f(t) = h(t)S(t)$. Selanjutnya persamaan tersebut dapat disubstitusikan ke persamaan berikut,

$$\begin{aligned} L(\beta, W, \sigma; t, x, \gamma) &= \prod_{i=1}^I \prod_{j=1}^{n_i} f_{ij}(t_{ij})^{\gamma_{ij}} S_{ij}(t_{ij})^{1-\gamma_{ij}} \\ &= \prod_{i=1}^I \prod_{j=1}^{n_i} [h(t_{ij}) S_{ij}(t_{ij})]^{\gamma_{ij}} S_{ij}(t_{ij})^{1-\gamma_{ij}} \\ &= \prod_{i=1}^I \prod_{j=1}^{n_i} h(t_{ij})^{\gamma_{ij}} S_{ij}(t_{ij})^{\gamma_{ij}} S_{ij}(t_{ij})^{1-\gamma_{ij}} \\ &= \prod_{i=1}^I \prod_{j=1}^{n_i} h(t_{ij})^{\gamma_{ij}} S_{ij}(t_{ij}) \end{aligned} \quad (4.22)$$

Sedangkan untuk parameter σ , β_i , dan λ memiliki distribusi posterior marginal dengan cara mengintergralkan keluar parameter-parameter yang bersangkutan dapat dijabarkan sebagai berikut,

$$p(\sigma | \lambda, \beta_i) \propto \int_{\lambda} \int_{\beta_1} \cdots \int_{\beta_p} l(t | \lambda, \beta_1, \dots, \beta_p) p(\lambda) p(\beta_1) \dots p(\beta_p) d\lambda d\beta_1, \dots, d\beta_p$$

$$p(\lambda | \sigma, \beta_i) \propto \int_{\sigma} \int_{\beta_1} \cdots \int_{\beta_p} l(t | \sigma, \beta_1, \dots, \beta_p) p(\sigma) p(\beta_1) \dots p(\beta_p) d\sigma d\beta_1, \dots, d\beta_p$$

$$\begin{aligned} p(\beta_1 | \sigma, \lambda, \beta_{i \neq 1}) &\propto \int_{\sigma} \int_{\beta_2} \cdots \int_{\beta_p} l(t | \sigma, \lambda, \beta_2, \dots, \beta_p) p(\sigma) p(\lambda) \\ &\quad \times p(\beta_2) \dots p(\beta_p) d\sigma d\lambda d\beta_2, \dots, d\beta_p \end{aligned}$$

$$p(\beta_p | \sigma, \lambda, \beta_{i \neq p}) \propto \int \int \cdots \int_a \int_{\beta_2} \int_{\beta_p} l(t | \sigma, \lambda, \beta_1, \dots, \beta_{p-1}) p(\sigma) p(\lambda) \\ \times p(\beta_1) \dots p(\beta_{p-1}) d\sigma d\lambda d\beta_1, \dots, d\beta_{p-1}$$

4.1.4. Estimasi Parameter Model Mixture Survival Spasial Dengan Menggunakan RJMCMC dan Gibbs Sampling

Pada bab 3 di penelitian ini sudah dijelaskan bahwa dilakukan pengujian dan penentuan distribusi dari data angka lama sekolah anak umu 16-18 tahun yang akan digunakan untuk model *survival* dan model *hazard* nya, dan jika terdapat indikasi bahwa distribusi dari data angka lama sekolah anak umur 16-18 tahun terdapat *mixture* dan cukup sulit untuk menentukan komponen *mixture*nya maka akan digunakan algoritma *Reversible jump Markov Chain Monte Carlo* (RJMCMC).

Keuntungan dari penggunaan metode ini adalah dapat memodelkan suatu *mixture* secara bersama-sama dengan banyaknya komponen penyusun *mixture* dan algoritma ini akan memudahkan untuk mendapatkan model *mixture* yang tepat bagi data yang terindikasi bersifat multimodal. Algoritma ini memanfaatkan proses *birth/death* dan *split/merge* dalam aplikasinya seperti yang tertulis pada algoritma pada Bab 2. Sedangkan untuk menentukan parameter pada model dengan menggunakan MCMC dengan *Gibbs Sampling*.

4.2. Model Mixture Survival Spasial Pada Angka Lama Sekolah Anak Umur 16-18 Tahun di Jawa Timur Berdasarkan Variabel-Variabel Yang Mempengaruhi.

Dalam sub bab ini akan dicari model *mixture survival* spasial pada angka lama sekolah anak umur 16-18 tahun di Jawa Timur berdasarkan variabel-variabel yang mempengaruhi tinggi atau rendahnya angka lama sekolah anak umur 16-18 tahun. Namun sebelumnya akan dibahas karakteristik responden anak umur 16-18 tahun berdasarkan variabel-variabel prediktornya.

4.2.1. Karakteristik Responden Anak Umur 16-18 Tahun Berdasarkan Variabel-Variabel Prediktornya

Pada sub bab ini dilakukan analisis statistika deskriptif untuk mengetahui karakteristik responden anak yang berumur 16-18 tahun. Analisis deskriptif ini dilakukan terhadap masing-masing variabel dalam penelitian meliputi variabel respon yaitu angka lama sekolah dan 8 variabel prediktor antara lain, variabel jenis kelamin (X_1), status perkawinan (X_2), status bekerja (X_3), tingkat pendidikan kepala rumah tangga (X_4), jenis kelamin kepala rumah tangga (X_5), banyaknya anggota rumah tangga (X_6), rata-rata pengeluaran rumah tangga per kapita (X_7), dan lokasi tempat tinggal (X_8). Untuk variabel yang berskala rasio akan dilihat nilai *mean*, varian, minimum dan maksimum dari variabel untuk mendapatkan informasi tentang variabel tersebut, sedangkan untuk variabel yang berskala nominal maupun ordinal untuk mendapatkan informasi tentang variabel tersebut akan dilihat dalam bentuk diagram batang dan *pie chart*. Data angka lama sekolah anak umur 16-18 tahun yang diolah dapat dilihat pada Lampiran 1.

Tabel 4.1. Deskriptif Responden Anak Berumur 16-18 Tahun

Variabel	Minimum	Maksimum	Mean	Std. Deviation
Angka Lama Sekolah Anak Umur 16-18 Tahun	0	12	9.39	1.841
Jumlah Anggota Rumah Tangga	1	11	4.41	1.342
Rata-Rata Pengeluaran Rumah Tangga	120955,94	13113030,12	517887,26	481380,3477

Sumber : pengolahan SPSS

Pada Tabel 4.1 didapatkan informasi bahwa waktu *survival* responden atau angka lama sekolah anak umur 16-18 tahun paling kecil 0 tahun yang artinya belum pernah menyentuh pendidikan formal. Paling besar yaitu 12 tahun yang artinya sudah mendapatkan ijazah pada jenjang SMA/ sederajat, sedangkan mean atau rata-rata nya 9,39 yang artinya sebagian besar anak yang berumur 16-18 tahun memiliki rata-rata lama sekolah di atas 9 tahun atau sudah lulus pada jenjang SMP/ sederajat. Untuk variabel jumlah anggota keluarga, paling sedikit ada rumah tangga yang memiliki 1 anggota keluarga dan yang paling banyak

hingga 11 anggota keluarga di dalamnya. Sedangkan untuk variabel rata-rata pengeluaran rumah tangga per kapita per bulan, pengeluaran terendah yaitu Rp.120.955,94 dan pengeluaran tertinggi hingga Rp.13.113.030,12, sedangkan untuk rata-rata dari rata-rata pengeluaran rumah tangga yaitu Rp.517.887,26.

Tabel 4.2. Deskriptif Status Sensor Anak Umur 16-18 Tahun Per Kabupaten/Kota Se Jawa Timur.

Kabupaten/Kota	Tersensor	%	Tidak Tersensor	%	Total
Pacitan	15	19%	66	81%	81
Ponorogo	10	12%	76	88%	86
Trenggalek	21	24%	66	76%	87
Tulungagung	16	14%	98	86%	114
Blitar	15	19%	62	81%	77
Kediri	28	23%	94	77%	122
Malang	34	25%	103	75%	137
Lumajang	24	29%	60	71%	84
Jember	30	24%	97	76%	127
Banyuwangi	22	18%	97	82%	119
Bondowoso	25	32%	54	68%	79
Situbondo	19	26%	54	74%	73
Probolinggo	23	29%	55	71%	78
Pasuruan	39	30%	89	70%	128
Sidoarjo	8	6%	120	94%	128
Mojokerto	14	9%	138	91%	152
Jombang	16	13%	110	87%	126
Nganjuk	14	14%	86	86%	100
Madiun	12	13%	82	87%	94
Magetan	8	12%	57	88%	65
Ngawi	8	9%	81	91%	89
Bojonegoro	21	17%	105	83%	126
Tuban	14	13%	97	87%	111
Lamongan	7	8%	86	92%	93
Gresik	8	6%	128	94%	136
Bangkalan	71	53%	63	47%	134
Sampang	60	58%	43	42%	103
Pamekasan	30	26%	86	74%	116
Sumenep	23	27%	63	73%	86
Kota Kediri	9	11%	75	89%	84
Kota Blitar	4	7%	55	93%	59
Kota Malang	13	13%	86	87%	99

Sumber : Pengolahan SPSS

Lanjutan Tabel 4.2

Kabupaten/Kota	Tersensor	%	Tidak Tersensor	%	Total
Kota Probolinggo	9	11%	73	89%	82
Kota Pasuruan	9	11%	73	89%	82
Kota Mojokerto	6	9%	61	91%	67
Kota Madiun	9	14%	54	86%	63
Kota Surabaya	19	12%	136	88%	155
Kota Batu	5	10%	47	90%	52
Jawa Timur	718	19%	3076	81%	3794

Deskriptif untuk angka lama sekolah anak umur 16-18 tahun di Jawa Timur dengan pengkategorian menjadi tersensor dan tidak tersensor dapat dilihat pada Tabel 4.2. bahwa untuk kategori tersensor dengan angka lama sekolah < 9 tahun, sedangkan untuk kategori tidak tersensor dengan rentang angka lama sekolah 9-12 tahun, maka didapatkan informasi bahwa jumlah anak yang memiliki angka lama sekolah 9-12 tahun atau dikatakan tidak tersensor pada Susenas 2012 sebanyak 3076 anak atau 81% dari jumlah anak yang berumur 16-18 tahun yang dipilih sebagai responden dalam Susenas 2012 sedangkan yang tidak tersensor dengan angka lama sekolah < 9 tahun atau dikatakan tidak lulus SMP/ sederajat sesuai program pemerintah dalam hal ini disebut tersensor sebanyak 718 anak atau 19% dari jumlah anak yang berumur 16-18 tahun yang dipilih sebagai responden dalam Susenas 2012. Angka lama sekolah anak umur 16-18 tahun di Jawa Timur berdasarkan kabupaten/kota didapatkan jumlah rumah tangga yang memiliki anak dengan umur 16-18 tahun disetiap kabupaten/kota beragam jumlahnya. Kabupaten/kota dengan jumlah rumah tangga yang terpilih sebagai responden dan memiliki anak umur 16-18 paling banyak yaitu kota Surabaya dengan 155 rumah tangga, sedangkan yang paling sedikit yaitu kota Batu dengan 52 rumah tangga.

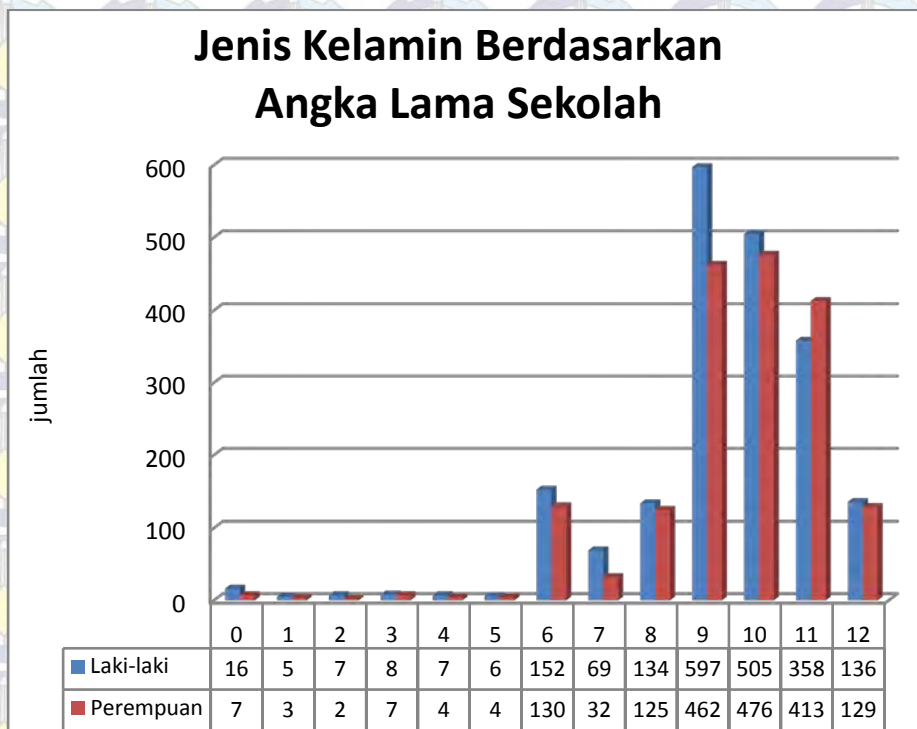
Jika melihat persentase dari jumlah anak yang tersensor dan yang tidak tersensor, yang tersensor paling banyak kabupaten Sampang dengan 60 anak atau 58% dari jumlah anak yang berumur 16-18 tahun yang dipilih sebagai responden di kabupaten Sampang tidak lulus pada jenjang SMP/ sederajat dan yang tersedikit kota Batu dengan 5 anak tetapi jika dilihat persentase yang tersedikit yaitu kabupaten Sidoarjo dengan 8 anak atau 6% dari jumlah anak yang berumur 16-18 tahun yang dipilih sebagai responden di kabupaten Sidoarjo tidak lulus pada

jenjang SMP/ sederajat. Sedangkan untuk tidak tersensor, paling banyak secara persentase yaitu kabupaten Gresik dengan 94 % (128 anak) dari jumlah anak yang berumur 16-18 tahun yang dipilih sebagai responden di kabupaten Gresik lulus pada jenjang SMP/ sederajat. Untuk yang tersedikit secara persentase yaitu kabupaten Sampang dengan 42% (43 anak) dari jumlah anak yang berumur 16-18 tahun yang dipilih sebagai responden di kabupaten Sampang lulus pada jenjang SMP/ sederajat.

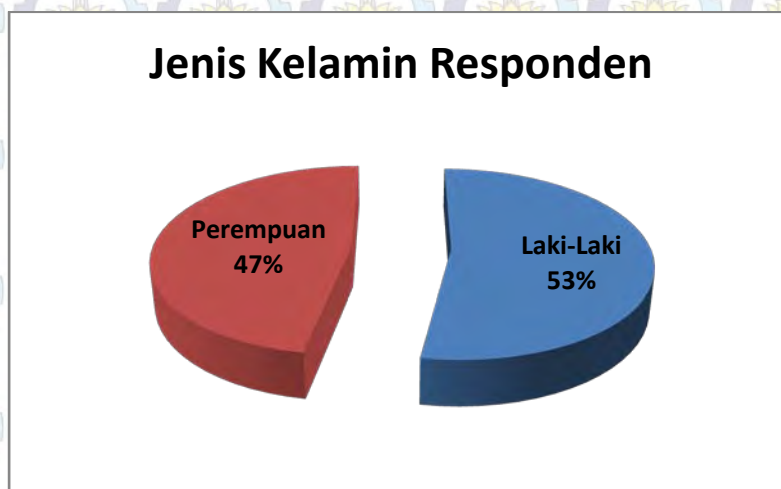
Seperti yang telah dijelaskan pada awal subbab ini, deskriptif dari variabel prediktor yang berkategori nominal dan ordinal akan dijelaskan dengan secara visual dengan menggunakan digram batang dan pie chart. Variabel-variabel yang akan dijelaskan dengan visual adalah variabel jenis kelamin responden, status perkawinan responden, status bekerja responden, jenis kelamin kepala rumah tangga, tingkat pendidikan terakhir kepala rumah tangga, dan lokasi tempat tinggal responden.

Jumlah responden anak laki-laki umur 16-18 tahun lebih banyak dari responden perempuan, hal tersebut dapat dilihat pada Gambar 4.2, jumlah responden laki-laki mencapai 53% dari jumlah responden anak umur 16-18 tahun secara keseluruhan sedangkan untuk responden perempuan hanya 47%. Sedangkan untuk perbedaan jumlah responden laki-laki dan perempuan disetiap angka lama sekolah dapat dilihat pada Gambar 4.1, bahwa disetiap jenjang angka lama sekolah responden laki-laki lebih banyak dari responden perempuan, kecuali pada angka lama sekolah 11 dimana jumlah responden perempuan lebih banyak.

Pada Gambar 4.1 juga dapat dilihat bahwa jumlah responden cenderung lebih banyak pada jenjang angka lama sekolah 9-12 tahun baik untuk responden laki-laki maupun responden perempuan.

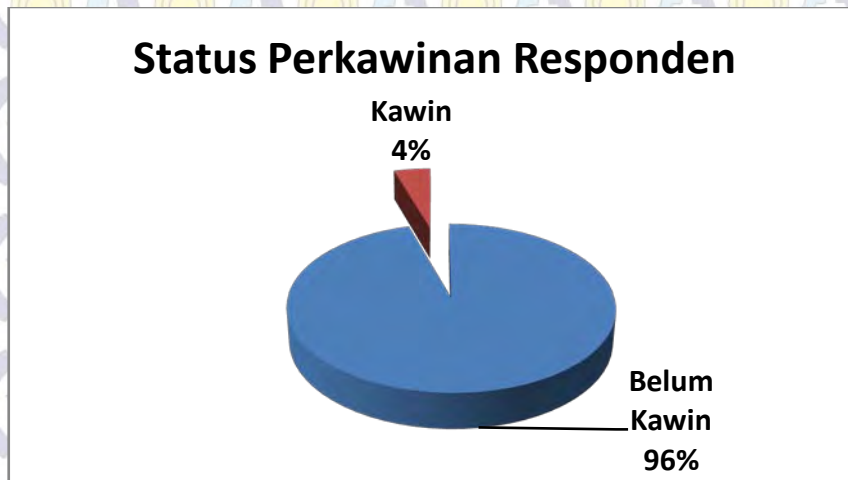


Gambar 4.1. Diagram Batang Variabel Jenis Kelamin Berdasarkan Angka Lama Sekolah Anak Umur 16-18 Tahun

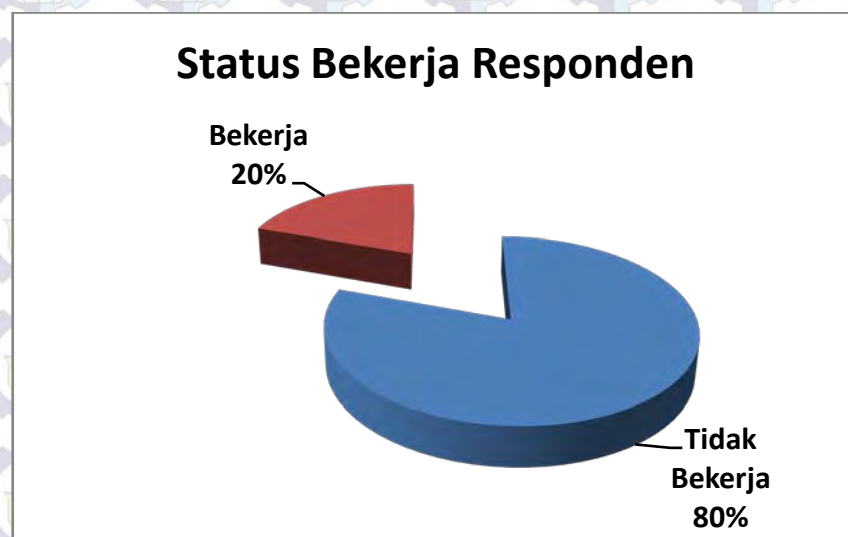


Gambar 4.2. Pie Chart Jenis Kelamin Responden

Pada variabel status perkawinan dari responden, terdapat 4% (sekitar 152 anak) responden yang sudah kawin dan sebagian besar dengan 96% belum kawin. Untuk status bekerja responden terdapat 20% responden yang sudah bekerja pada umur 16-18 tahun sedangkan sisanya 80% tidak bekerja.



Gambar 4.3. Pie Chart Status Perkawinan dari Responden



Gambar 4.4. Pie Chart Status Bekerja dari Responden

Berdasarkan Tabel 4.3, dapat dilihat bahwa responden anak umur 16-18 tahun dengan status bekerja lebih banyak pada rumah tangga dengan rata-rata pengeluaran <Rp. 500.000 yaitu 594 anak, sedangkan untuk anak dengan penghasilan rumah tangga pada rentang Rp. 500.001 – Rp. 1.000.000 hanya 130 anak dan sisanya 18 anak dengan status bekerja pada rumah tangga dengan pengeluaran > Rp. 1.000.000. Untuk status responden anak yang tidak bekerja lebih banyak dari yang bekerja, hal tersebut mungkin dikarenakan orangtua

mereka tidak memperkenankan anak mereka untuk bekerja dahulu pada usia sekolah.

Tabel 4.3. Deskriptif Rata-Rata Pengeluaran terhadap Status Bekerja Responden

Rata-Rata Pengeluaran	Status_Bekerja		Total
	Tidak	Ya	
<Rp. 500.000	1887	594	2481
Rp. 500.001 - Rp. 1000.000	912	130	1042
>Rp. 1000.001	253	18	271
Total	3052	742	3794

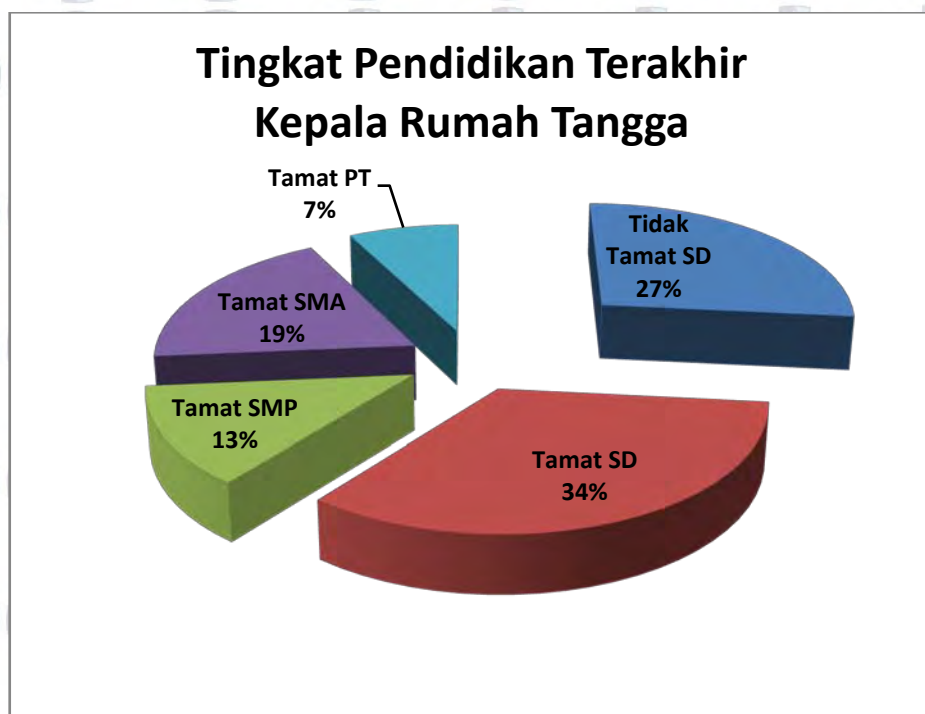
Dari penelitian sebelumnya tentang angka lama sekolah yang terdahulu, jenis kelamin kepala rumah tangga mempengaruhi angka lama sekolah dari anak-anak mereka. Pada penelitian ini kepala rumah tangga berjenis kelamin laki-laki lebih banyak dari kepala rumah tangga berjenis kelamin perempuan, yaitu 90% (3410 kepala rumah tangga) yang berjenis kelamin laki-laki dan 10% (384 kepala rumah tangga) yang berjenis kelamin perempuan.



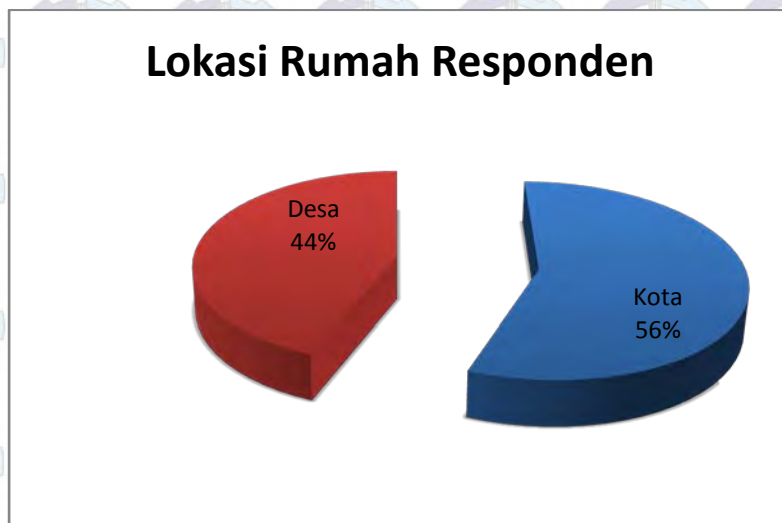
Gambar 4.5. Pie Chart Jenis Kelamin Kepala Rumah Tangga Responden

Sedangkan untuk tingkat pendidikan terakhir kepala rumah tangga, bahwa 27 % (1009 kepala rumah tangga) tidak tamat SD/ sederajat atau dapat dikatakan tidak mendapat pendidikan formal, 34% (1293 kepala rumah tangga) hanya memiliki ijazah sampai jenjang SD/ sederajat, 13% (490 kepala rumah tangga) memiliki ijazah sampai jenjang SMP/ sederajat, 19% (714 kepala rumah tangga) memiliki ijazah sampai jenjang SMA/ sederajat dan sisanya 7% (288 kepala rumah tangga) memiliki ijazah sampai perguruan tinggi.

Untuk variabel lokasi tempat tinggal responden, 56% dari total responden atau 2133 responden yang terpilih bertempat tinggal di perkotaan, sedangkan sisanya 44% dari total responden atau 1661 responden yang terpilih bertempat tinggal di perdesaan.



Gambar 4.6. Pie Chart Tingkat Pendidikan Terakhir Kepala Rumah Tangga Responden

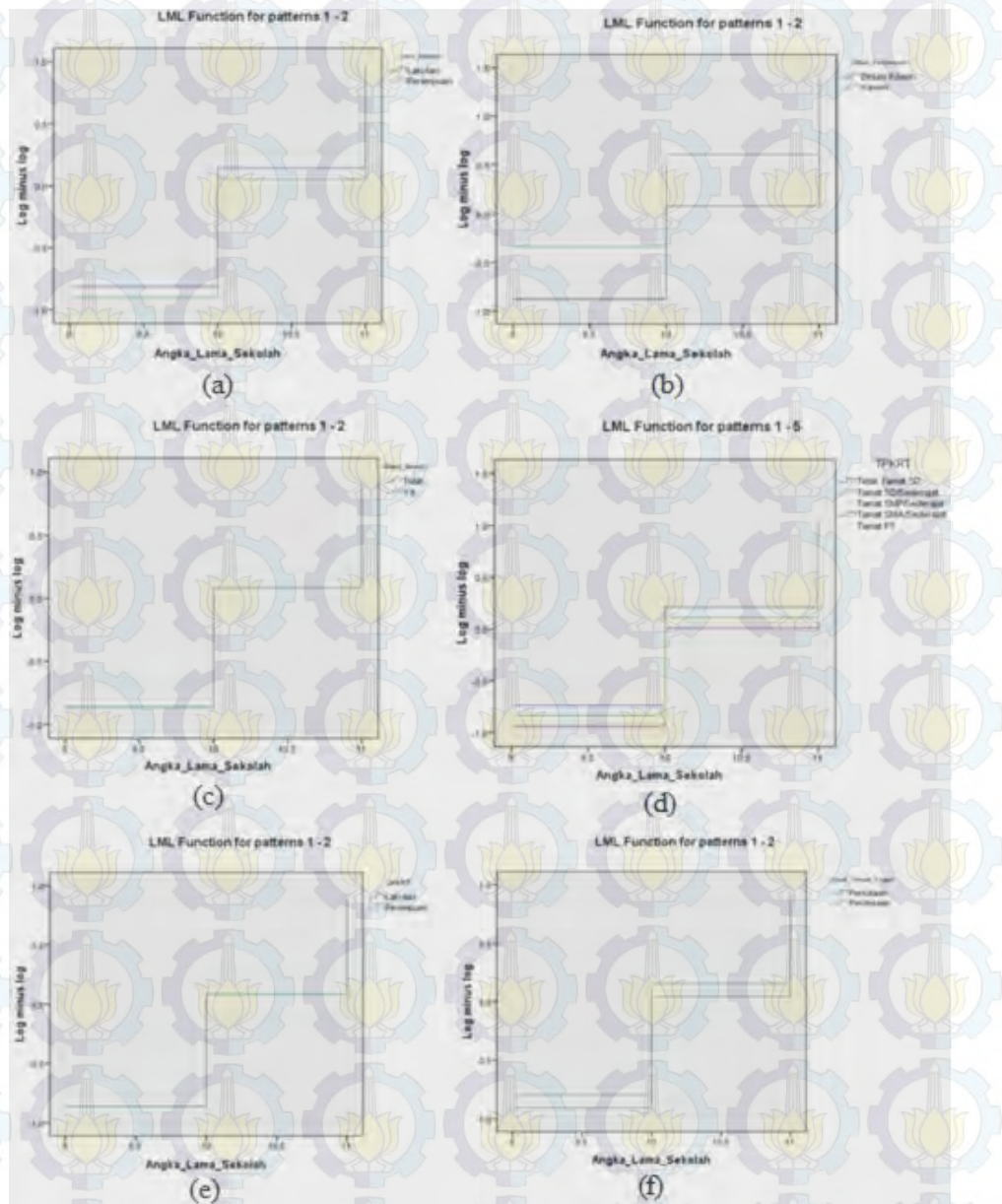


Gambar 4.7. Pie Chart Lokasi Rumah Responden

4.2.2. Pengujian Asumsi Pemodelan *Hazard Proporsional*

Asumsi pemodelan yang harus dipenuhi dalam *hazard* proporsional adalah bahwa fungsi *hazard* dari variabel prediktor yang bersifat kategorik harus proporsional setiap waktu. Pengujian dapat dilakukan dengan menggunakan plot $\ln[-\ln S(t)]$. Asumsi *hazard* proporsional terpenuhi bila garis antar kategori sejajar.

Telah dijelaskan pada bab II bahwa pada penelitian ini terdapat enam variabel yang bersifat kategorik dari delapan variabel. Dari Gambar 4.8 (a), (b), (c), (d), (e), dan (f), dapat dilihat bahwa garis antar kategori sejajar sehingga bisa dikatakan asumsi *hazard* proporsional sudah terpenuhi untuk keenam variabel kategorik, artinya variabel prediktor telah independen terhadap waktu serta hubungan antara kumulatif *hazard* sudah proporsional/ konstan setiap waktu.



Gambar 4.8. Plots $\ln[-\ln S(t)]$ Untuk Uji Proporsional Hazard (a). Jenis Kelamin Responden, (b). Status Perkawinan Responden, (c). Status Bekerja Responden, (d). Tingkat Pendidikan Terakhir Kepala Rumah Tangga, (e). Jenis Kelamin Kepala Rumah Tangga, dan (f). Lokasi Tempat Tinggal Responden.

4.2.1. Penimbang Spasial

Pendekatan yang digunakan dalam menentukan penimbang spasial dalam penelitian ini menggunakan pendekatan area. Penimbang spasial yang diperoleh selanjutnya merupakan salah satu parameter dalam prior CAR yang menjadi distribusi dari efek random (*frailty*) model *survival*. Jenis penimbang yang

digunakan yaitu penimbang spasial *Queen Contiguity* (persinggungan sisi dan sudut) dan kabupaten yang saling bertetangga disebut sebagai *adjacent* sehingga ketetanggaan dari masing-masing kabupaten di provinsi Jawa Timur akan disusun dalam matriks *adjacent*.

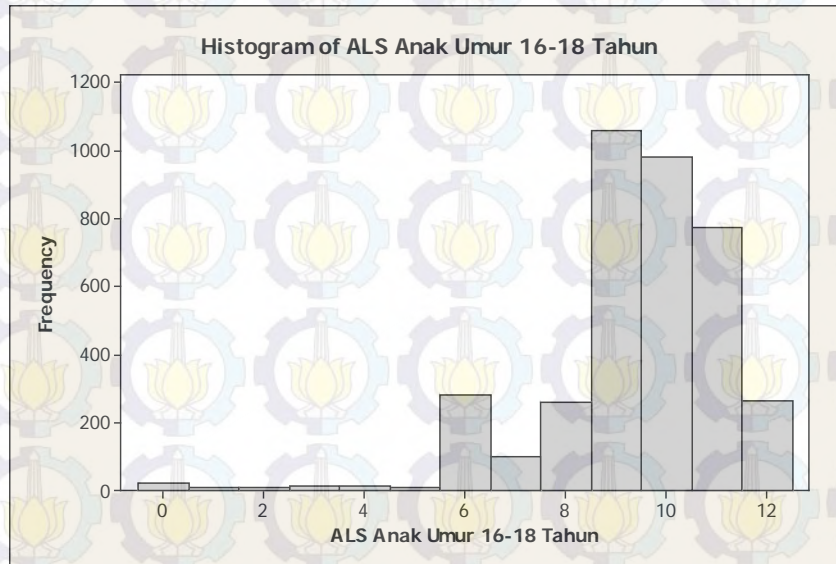
Langkah dalam menentukan matriks *adjacent* dengan mengolah peta/map poligon Jawa Timur per kabupaten yang sudah didapat melalui *software* GeoDa. Hasil dari penimbang spasial *Queen Contiguity* yang telah didapat akan berupa matriks dengan ukuran 38×38 (dikarenakan banyak kabupaten/kota di Jawa Timur sebanyak 38 kabupaten/kota) yang element dari matriks tersebut 0 dan 1, 1 untuk daerah yang bertetangga dengan daerah j (daerah yang dilihat ketetanggannya), sedangkan 0 untuk daerah yang tidak bertetangga dengan daerah j . Lalu matriks ini distandarisasi dengan cara membagi setiap barisnya dengan banyak daerah yang bertetangga dengan daerah j . Hasil matriks yang sudah distandarisasi dapat dilihat pada lampiran. Secara keseluruhan matriks *adjacent* ini memiliki 140 ketetanggaan dari 38 kabupaten yang diteliti.

4.2.2. Mengidentifikasi dan Menguji Distribusi Data Waktu *Survival* Angka Lama Sekolah

Pendugaan distribusi dilakukan terhadap waktu *survival* (t) dimana pada penelitian ini adalah angka lama sekolah anak umur 16-18 tahun di Jawa Timur. Gambar 4.9 menunjukkan histogram dari waktu *survival* responden anak umur 16-18 tahun, dapat dilihat bahwa diindikasikan terdapat lebih dari satu puncak sehingga diduga terdapat kasus multimodal pada waktu *survival* ini sehingga perlu digunakan distribusi *mixture* agar hasil estimasi distribusi dan hasil analisis selanjutnya bisa lebih akurat, hal tersebut didukung dengan hasil pengujian Anderson-Darling pada Lampiran 2 dengan tidak ditemukan distribusi yang sesuai untuk data angka lama sekolah anak umur 16-18 tahun.

Tidak semua distribusi *mixture* bisa terbaca atau dapat ditentukan berapa banyak komponennya hanya dengan gambaran visual dari histogram saja. Metode menentukan banyak komponen *mixture* dengan melihat visual dari histogram disebut metode prediktif, dimana hanya memprediksi berapa banyak komponen *mixture* berdasarkan banyak puncak dari data. Untuk itu diperlukan suatu metode dimana dimana penentuan jumlah komponen *mixture*-nya lebih akurat dan dapat

dipercaya, untuk itu dalam penelitian ini untuk penentuan banyak komponen distribusi *mixture* agar lebih akurat digunakan metode *reversible jump* markov chain monte carlo.



Gambar 4.9. Histogram Angka Lama Sekolah Anak Umur 16-18 Tahun

Dengan hipotesis bahwa

H0 : Waktu *survival* sesuai dengan distribusi dugaan (Lognormal/Normal/Weibull/Log Logistik)

H1 : Waktu *survival* tidak sesuai dengan distribusi dugaan (Lognormal/Normal/Weibull/Log Logistik)

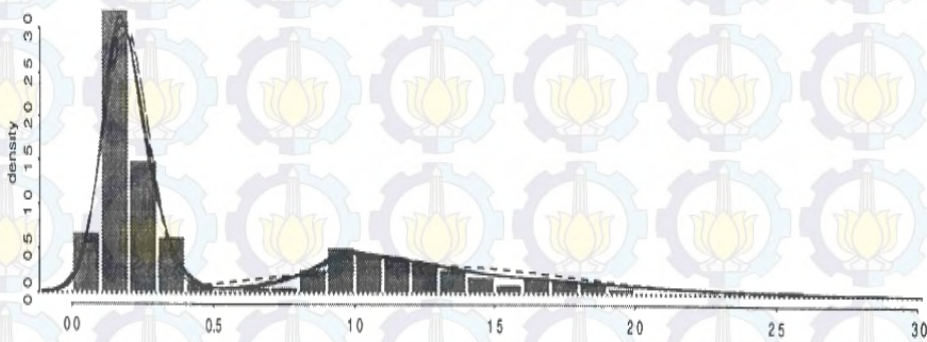
Statistik uji sesuai dengan persamaan (2.1) dan daerah kritis didasarkan pada kondisi tolak H₀ jika $A_n^2 > a_{n,1-\alpha}$ atau P-value $< \alpha$, dengan $\alpha = 0,05$ maka hasilnya dapat dilihat pada Tabel 4.4 dan Lampiran 2. Distribusi yang paling mendekati untuk data angka lama sekolah anak umur 16-18 tahun yaitu distribusi Normal 2-parameter. Distribusi Normal 2-parameter ini akan dijadikan distribusi utama dalam penerapan *reversible jump* markov chain monte carlo.

Tabel 4.4. Uji Distribusi Waktu *Survival*

Distribusi	Statistik Uji (A_n^2)	Nilai Kritis ($\alpha_{n,1-\alpha}$)	Keputusan
Log Normal	281,78	2,5018	Tolak H_0
Normal	146,06	2,5018	Tolak H_0
Weibull	147,00	2,5018	Tolak H_0
Log Logistik	193,50	2,5018	Tolak H_0

Sumber : Program Easy Fit

Distribusi Normal adalah salah satu distribusi non-negatif. sedangkan jika dilihat berdasarkan Gambar histogram (4.9), tidak ada angka yang bernilai negatif. Pendekatan dengan distribusi normal yang dimaksud untuk mendekati setiap bagian komponen *mixture* dengan distribusi normal.

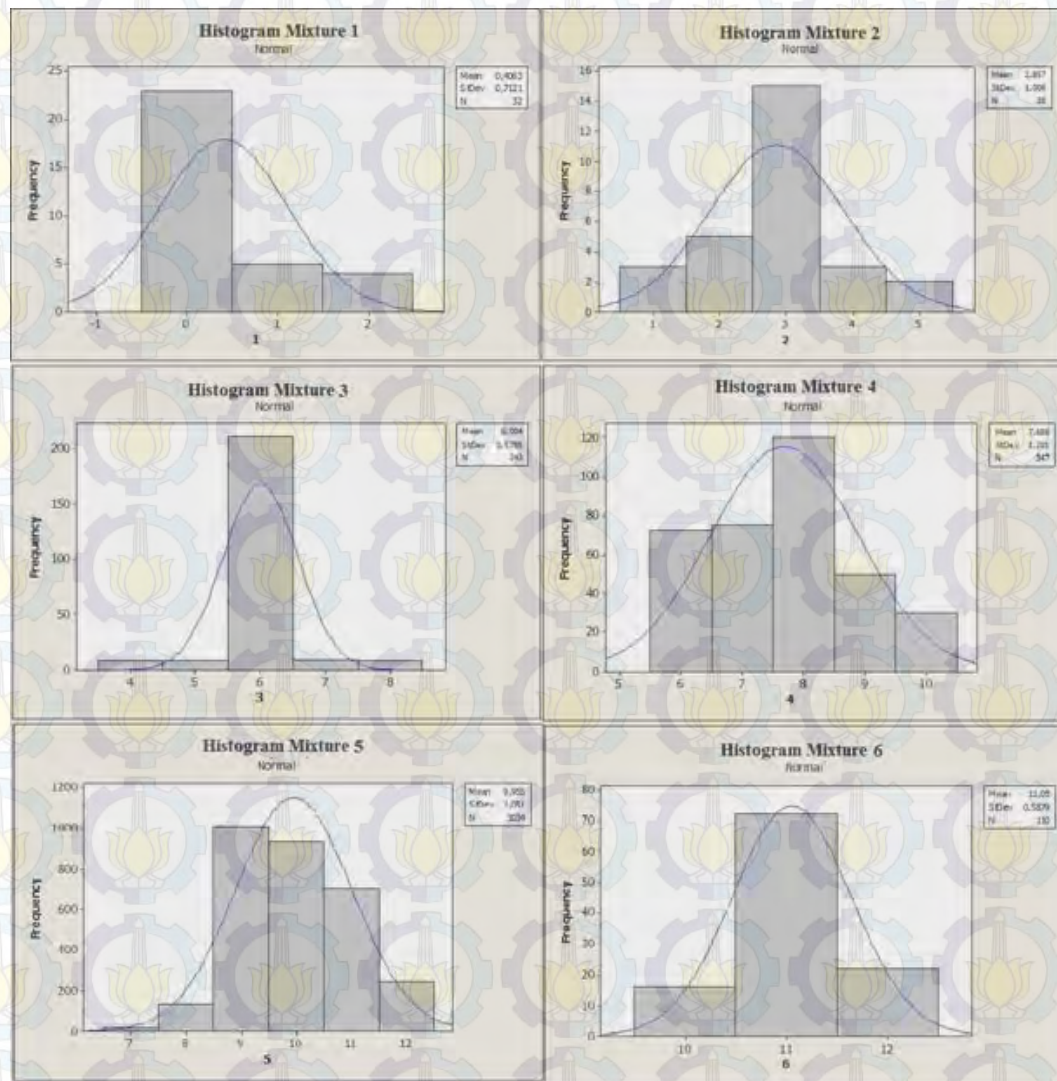


Gambar 4.10. Contoh Pendekatan Distribusi Normal Pada Komponen *Mixture* Pada Penelitian Green dan Sylvia (1997) Data Enzyme

4.2.3. Menentukan Model dan Parameter Model *Mixture Survival* Spasial

Sebelum menentukan model *mixture survival* spasial akan ditentukan dahulu banyak komponen yang dibentuk oleh model dengan menggunakan algoritma RJMCMC, adapun syntax dan hasil dapat dilihat pada Lampiran 5. Berdasarkan hasil yang diperoleh, dengan menggunakan iterasi sebanyak 100000, didapatkan banyak komponen untuk distribusi *mixture* pada data *survival* angka lama sekolah anak umur 16-18 tahun yaitu sebanyak 6 komponen (lihat Gambar 4.11). Hal tersebut dapat dilihat bahwa hanya 6 komponen yang memiliki nilai posterior probability yang menjadi nilai komponen awal setiap *mixture*, yaitu sebesar 0,86021, 0,05878, 0,05303, 0,01396, 0,00557, dan 0,00845. Nilai

komponen awal ini digunakan untuk menentukan data yang sesuai dengan komponennya.



Gambar 4.11. Histogram Data *Mixture* 1 Sampai *Mixture* 6

Jumlah komponen *mixture* yang telah diketahui digunakan untuk membentuk model *mixture survival* spasial dengan variabel-variabel yang mempengaruhi. Digunakan software WinBUGS untuk membantu dalam perhitungan serta penentuan nilai parameter-parameter pada persamaan 4.20.

Tabel 4.5. Estimasi Parameter Model *Mixture Survival Spasial*

Node	Mean	Sd	MC Error	2.5%	Median	97.5%
Phi[1]	0.00869	0.001481	0,0332	0.005987	0.0086	0.0118
Phi[2]	0,00767	0,001393	0,0279	0,005288	0,007587	0,01066
Phi[3]	0,0642	0,004079	0,0913	0,05663	0,06416	0,07238
Phi[4]	0,09174	0,004648	0,00008	0,08291	0,0916	0,101
Phi[5]	0,7984	0,006549	0,133	0,7856	0,7986	0,8113
Phi[6]	0,02929	0,002773	0,0561	0,02411	0,02919	0,03501
b1[1]	0,02705	0,08932	0,01272	-0,1669	0,02487	0,1964
b2[1]	0,9876	0,1508	0,02214	0,7003	0,9863	1,288
b3[1]	1,328	0,1621	0,02415	1,091	1,32	1,599
b4_1[1]	-0,08707	0,1801	0,02625	-0,5033	-0,07932	0,2023
b4_2[1]	0,07857	0,3273	0,04377	-0,6468	0,07265	0,7583
b4_3[1]	-0,1805	0,2108	0,03023	-0,5683	-0,1531	0,1717
b4_4[1]	-0,08752	0,3676	0,0499	-0,7417	-0,1116	0,7962
b5[1]	0,4103	0,09288	0,0131	0,2095	0,408	0,5642
b6[1]	0,003723	0,02156	0,002909	-0,03822	8,45E-01	0,05784
b7[1]	-0,02557	0,02023	0,002779	-0,07475	-0,02323	0,003989
b8[1]	-0,8775	0,08503	0,01228	-1,015	-0,8855	-0,6988
b1[2]	0,5267	0,02618	0,003912	0,479	0,5311	0,5612
b2[2]	0,3795	0,01732	0,002565	0,3456	0,3817	0,4089
b3[2]	-1,028	0,0147	0,002109	-1,051	-1,032	-0,9999
b4_1[2]	0,02426	0,07201	0,0107	-0,09264	0,04424	0,1283
b4_2[2]	0,6229	0,0595	0,008888	0,5324	0,613	0,7132
b4_3[2]	-1,329	0,2244	0,02558	-1,825	-1,32	-0,8784
b4_4[2]	-1,766	0,2069	0,02584	-2,241	-1,768	-1,43
b5[2]	-0,8042	0,02204	0,003236	-0,8425	-0,803	-0,769
b6[2]	0,1589	0,005748	0,844	0,1485	0,16	0,1685
b7[2]	0,7912	0,002474	0,361	0,7861	0,7922	0,7945
b8[2]	0,5496	0,00953	0,001406	0,5362	0,5468	0,5681
b1[3]	0,5389	0,01158	0,001692	0,5177	0,538	0,5625
b2[3]	0,2462	0,01038	0,001531	0,2274	0,2447	0,2639
b3[3]	0,1697	0,01448	0,002136	0,1403	0,1742	0,1933
b4_1[3]	0,01034	0,01926	0,002817	-0,01658	0,004454	0,04662
b4_2[3]	0,2694	0,02044	0,003007	0,23	0,2701	0,301
b4_3[3]	-1,024	0,1288	0,01812	-1,266	-1,021	-0,7419
b4_4[3]	0,1204	0,3459	0,05139	-0,5799	0,2927	0,5987
b5[3]	-0,2245	0,01666	0,00247	-0,2509	-0,2278	-0,1948

Lanjutan 1 Tabel 4.5.

Node	Mean	Sd	MC Error	2.5%	Median	97.5%
b6[3]	0,5691	0,002693	0,401	0,5652	0,5688	0,5729
b7[3]	0,3746	0,003618	0,539	0,3694	0,3743	0,3811
b8[3]	0,01988	0,004412	0,632	0,01068	0,02018	0,02699
b1[4]	0,09722	0,03892	0,005682	0,02653	0,09561	0,1833
b2[4]	0,346	0,05114	0,00753	0,2371	0,3584	0,4167
b3[4]	0,05621	0,03718	0,005375	0,001356	0,04833	0,1394
b4_1[4]	1,294	0,1004	0,01461	1,136	1,272	1,492
b4_2[4]	1,262	0,1105	0,01595	1,089	1,232	1,515
b4_3[4]	1,375	0,1336	0,01912	1,149	1,383	1,617
b4_4[4]	1,302	0,1066	0,01574	1,103	1,298	1,514
b5[4]	0,1496	0,08143	0,01203	0,0139	0,1631	0,2691
b6[4]	0,07035	0,002363	0,350	0,06517	0,07105	0,07355
b7[4]	0,02315	0,01196	0,001731	0,001777	0,02325	0,04526
b8[4]	0,0706	0,03307	0,004793	0,01068	0,06714	0,1427
b1[5]	0,1418	0,01107	0,001421	0,1194	0,1418	0,1627
b2[5]	0,2348	0,01508	0,001908	0,203	0,237	0,2607
b3[5]	0,1695	0,009682	0,00108	0,1528	0,1687	0,1882
b4_1[5]	0,603	0,0316	0,004293	0,5353	0,602	0,6664
b4_2[5]	0,6197	0,03245	0,004465	0,5482	0,6193	0,691
b4_3[5]	0,67	0,0315	0,004173	0,6084	0,6691	0,7387
b4_4[5]	0,69	0,02999	0,003995	0,6328	0,69	0,7594
b5[5]	0,2619	0,0141	0,001771	0,2347	0,2604	0,2922
b6[5]	0,01332	5,71E-4	0,0393	0,01206	0,01334	0,01443
b7[5]	0,07189	0,003826	0,508	0,06466	0,07165	0,07988
b8[5]	0,2039	0,0107	0,001354	0,1821	0,2045	0,2215
b1[6]	0,1225	0,02813	0,003531	0,05828	0,124	0,1707
b2[6]	0,7286	0,09177	0,01305	0,5384	0,7199	0,912
b3[6]	0,2573	0,02996	0,003523	0,2015	0,2576	0,311
b4_1[6]	-0,04943	0,08181	0,01005	-0,2002	-0,04617	0,1035
b4_2[6]	0,153	0,06349	0,007874	0,03974	0,1538	0,2768
b4_3[6]	0,09768	0,06401	0,007685	-0,02007	0,09498	0,2263
b4_4[6]	0,05246	0,05544	0,006933	-0,04913	0,04924	0,1588
b5[6]	0,1959	0,03887	0,004955	0,1236	0,1953	0,2728
b6[6]	0,004926	0,002667	0,271	-8,16E-01	0,004963	0,009834
b7[6]	0,05403	0,0124	0,001632	0,02615	0,05466	0,0779
b8[6]	0,5175	0,062	0,008505	0,3849	0,5171	0,6355

Lanjutan 2 Tabel 4.5.

Node	Mean	Sd	MC Error	2.5%	Median	97.5%
ptau[1]	3,66	4,141	0,443	0,1324	2,147	15,8
ptau[2]	0,1962	0,406	0,04869	0,001191	0,04954	1,423
ptau[3]	0,01153	0,006376	0,0007346	0,003676	0,01007	0,02717
ptau[4]	0,2724	0,04062	0,002609	0,1983	0,2693	0,3602
ptau[5]	0,1394	0,003863	0,161	0,1317	0,1394	0,1468
ptau[6]	0,3749	0,05469	0,002521	0,2781	0,3719	0,4909
sigma	0,3562	0,04285	0,00135	0,2837	0,3522	0,4537
dtau	8,216	1,915	0,06194	4,863	8,061	12,45

Sumber : Pengolahan WinBUGS

Seperti yang telah dijelaskan pada bab 3, bahwa variabel-variabel yang diduga berpengaruh terhadap laju kelulusan anak umur 16-18 tahun meliputi jenis kelamin (X1), status perkawinan (X2), status bekerja (X3), tingkat pendidikan kepala rumah tangga (X4), jenis kelamin kepala rumah tangga (X5), rata-rata pengeluaran rumah tangga per kapita (X6), Jumlah anggota rumah tangga (X7), dan lokasi tempat tinggal (kota atau desa) (X8). Hasil estimasi (*posterior summaries*) terhadap variabel-variabel yang mempengaruhi laju kelulusan responden anak umur 16-18 tahun secara lengkap dapat dilihat pada Lampiran 6.

Hasil estimasi parameter (*posterior summaries*) model *mixture survival* spasial bedistribusi Normal disajikan pada Tabel 4.5. Variabel-variabel yang dianggap signifikan mempengaruhi laju kelulusan responden anak umur 16-18 tahun jika nilai pada selang 2,5% hingga 97,5% tidak memuat nilai 0. Tabel 4.5 menunjukkan bahwa tidak semua variabel berpengaruh signifikan terhadap laju kelulusan responden anak umur 16-18 tahun dan terdapat komponen *mixture* yang dipengaruhi oleh bebrapa variabel saja. Kolom *node* merupakan kovariat (variabel-variabel) yang diduga berpengaruh terhadap laju kelulusan responden anak umur 16-18 tahun, kolom *mean* menunjukkan besarnya nilai parameter model, sedangkan ketiga kolom berikutnya besarnya nilai estimasi pada selang kepercayaan 95%.

Pada komponen *mixture* pertama dengan besar proporsi ($\Phi[1]$) 0,00869, responden yang masuk yaitu responden dengan angka lama sekolah 0, 1 dan 2. Variabel-variabel yang secara signifikan mempengaruhi model meliputi status

perkawinan responden (X2), status bekerja responden (X3), jenis kelamin kepala rumah tangga (X5), dan lokasi tempat tinggal responden (X8). Ciri khas dari kelompok *mixture* pertama ini yaitu daerah tempat tinggal responden yang di daerah pedesaan.

Untuk komponen *mixture* kedua dengan besar proporsi ($\Phi[2]$) 0,00767, responden yang masuk yaitu responden dengan angka lama sekolah 1, 2, 3, 4, dan 5, serta variabel-variabel yang secara signifikan mempengaruhi model meliputi semua variabel kecuali variabel dummy kategori 1 pada variabel tingkat pendidikan tertinggi kepala keluarga. Ciri khas dari kelompok *mixture* kedua ini yaitu tingkat pendidikan kepala rumah tangga yang didominasi tidak tamat SD dan lokasi rumah pedesaan.

Untuk komponen *mixture* ketiga dengan besar proporsi ($\Phi[3]$) 0,0642, responden yang masuk yaitu responden dengan angka lama sekolah 4, 5, 6, 7, dan 8. Variabel-variabel yang secara signifikan mempengaruhi model meliputi semua variabel kecuali variabel dummy kategori 1 dan kategori 4 pada variabel tingkat pendidikan tertinggi kepala keluarga, dengan ciri khas dari kelompok *mixture* ketiga ini tingkat pendidikan kepala rumah tangga yang secara signifikan mempengaruhi dan mendominasi yaitu kepala rumah tangga dengan tingkat pendidikan tertinggi tamat SMP

Untuk komponen *mixture* keempat dengan besar proporsi ($\Phi[4]$) 0,09174, responden yang masuk yaitu responden dengan angka lama sekolah 6, 7, 8, 9, dan 10, dimana semua variabel secara signifikan mempengaruhi model pada *mixture* keempat. Ciri khas dari kelompok *mixture* keempat ini yaitu tingkat pendidikan tertinggi kepala rumah tangga yang mempunyai ijazah hampir seimbang dengan yang tidak punya ijazah (tidak tamat SD)

Untuk *mixture* kelima dengan besar proporsi ($\Phi[5]$) 0,7984, responden yang masuk yaitu responden dengan angka lama sekolah 8, 9, 10, 11, dan 12, dimana sama hal nya seperti *mixture* keempat dimana semua variabel secara signifikan mempengaruhi model pada *mixture* kelima. Ciri khas dari kelompok *mixture* kelima ini yaitu tempat tinggal responden yang didominasi pada daerah perkotaan.

Sedangkan untuk *mixture* keenam dengan besar proporsi (Phi[6]) 0,02929, responden yang masuk yaitu responden dengan angka lama sekolah 10, 11, dan 12. Variabel-variabel yang secara signifikan mempengaruhi model pada *mixture* keenam meliputi jenis kelamin responden (X1), status perkawinan responden (X2), status bekerja responden (X3), jenis kelamin kepala rumah tangga (X5), jumlah anggota rumah tangga (X7), dan lokasi tempat tinggal responden (X8). Ciri khas dari kelompok *mixture* keenam ini yaitu hampir semua responden dengan status belum kawin dan mendominasi tempat tinggal di daerah perkotaan serta tingkat pendidikan tertinggi kepala rumah tangga yang didominasi tamat SMA.

Dalam menginterpretasikan tingkat resiko (kecenderungan) suatu variabel digunakan nilai *odds ratio* yang merupakan suatu perbandingan individu pada kondisi variabel/variabel prediktor (x) kategori yang diharapkan dengan variabel/variabel prediktor (x) pada kategori pembandingan.

Tabel 4.6. Odds Ratio

Node	Odds Ratio	Node	Odds Ratio	Node	Odds Ratio
<i>Mixture 1</i>		<i>Mixture 2</i>		<i>Mixture 3</i>	
b2[1]	2,68	b1[2]	1,69	b1[3]	1,71
b3[1]	3,77	b2[2]	1,46	b2[3]	1,28
b5[1]	1,51	b3[2]	0,36	b3[3]	1,18
b8[1]	0,42	b4_2[2]	1,86	b4_2[3]	1,31
		b4_3[2]	0,26	b4_3[3]	0,36
		b4_4[2]	0,17	b5[3]	0,80
		b5[2]	0,45	b6[3]	1,77
		b6[2]	1,17	b7[3]	1,45
		b7[2]	2,21	b8[3]	1,02
		b8[2]	1,73		

Sumber : Pengolahan Excel

Lanjutan Tabel 4.6.

Node	Odds Ratio	Node	Odds Ratio	Node	Odds Ratio
<i>Mixture 4</i>		<i>Mixture 5</i>		<i>Mixture 6</i>	
b1[4]	1,10	b1[5]	1,15	b1[6]	1,13
b2[4]	1,41	b2[5]	1,26	b2[6]	2,07
b3[4]	1,06	b3[5]	1,18	b3[6]	1,29
b4_1[4]	3,65	b4_1[5]	1,83	b4_2[6]	1,17
b4_2[4]	3,53	b4_2[5]	1,86	b5[6]	1,22
b4_3[4]	3,96	b4_3[5]	1,95	b7[6]	1,06
b4_4[4]	3,68	b4_4[5]	1,99	b8[6]	1,68
b5[4]	1,16	b5[5]	1,30		
b6[4]	1,07	b6[5]	1,01		
b7[4]	1,02	b7[5]	1,07		
b8[4]	1,07	b8[5]	1,23		

Sumber : Pengolahan Excel

Berdasarkan Tabel 4.5 dan Tabel 4.6, variabel-variabel yang dinyatakan berpengaruh signifikan terhadap laju kelulusan responden anak umur 16-18 tahun dengan selang kepercayaan 95% beserta interperasinya berdasarkan nilai *odds ratio* sebagai berikut,

1. Variabel Jenis Kelamin Responden (X1)

Variabel jenis kelamin responden (X1) secara signifikan mempengaruhi model *mixture survival* spasial pada *mixture* kedua ($b1[2]=0,5627$), ketiga ($b1[3]=0,5389$), keempat ($b1[4]=0,0972$), kelima ($b1[5]=0,1418$) dan keenam ($b1[6]=0,1225$). Pada variabel jenis kelamin yang menjadi acuan untuk tingkat kecenderungan yaitu responden dengan kategori jenis kelamin laki-laki, sedangkan responden kategori jenis kelamin perempuan sebagai pembanding.

Pada *mixture* kedua, tingkat kecenderungan responden dengan jenis kelamin laki-laki untuk memperoleh kesempatan sekolah lebih lama sebesar 1,69 kali daripada responden dengan jenis kelamin perempuan, artinya jika ada responden perempuan dengan angka lama sekolah 7 tahun maka responden laki-laki memiliki angka lama sekolah minimal 12 tahun.

Pada *mixture* ketiga sebesar 1,71 kali, pada *mixture* keempat sebesar 1,10 kali, pada *mixture* kelima sebesar 1,15 kali, sedangkan pada *mixture* keenam, sebesar 1,13 kali.

2. Variabel Status Perkawinan Responden (X2)

Variabel status perkawinan responden (X2) secara signifikan mempengaruhi model *mixture survival* spasial pada semua komponen *mixture*. Dengan nilai parameter dari variabel status perkawinan responden pada *mixture* pertama (b2[1]) 0,9876, *mixture* kedua (b2[2]) 0,3795, *mixture* ketiga (b2[3]) 0,2462, *mixture* keempat (b2[4]) 0,346, *mixture* kelima (b2[5]) 0,2348 dan *mixture* keenam (b2[6]) 0,7286. Pada variabel status perkawinan yang menjadi acuan untuk tingkat kecenderungan yaitu responden dengan kategori status belum kawin, sedangkan responden kategori status sudah kawin sebagai pembanding.

Pada *mixture* pertama, tingkat kecenderungan responden yang berstatus belum kawin untuk memperoleh kesempatan sekolah lebih lama sebesar 2,68 kali daripada responden yang sudah kawin, artinya jika ada responden dengan status kawin putus sekolah pada angka lama sekolah 4 tahun maka responden yang belum kawin akan memiliki angka lama sekolah minimal 11 tahun. Hal tersebut dikarenakan jika seorang anak kawin pada jenjang pendidikan SD, SMP, atau SMA, anak tersebut dikeluarkan dari sekolah/*drop out*, maka responden tersebut sulit mendapatkan kesempatan sekolah lebih lama.

Pada *mixture* kedua sebesar 1,46 kali, pada *mixture* keempat sebesar 1,41 kali, pada *mixture* kelima sebesar 1,27 kali, sedangkan pada *mixture* keenam sebesar 2,07.

3. Variabel Status Bekerja Responden (X3)

Variabel status bekerja responden (X3) secara signifikan mempengaruhi model *mixture survival* spasial pada semua komponen *mixture*. Dengan nilai parameter dari variabel status bekerja responden pada *mixture* pertama (b3[1]) 1,328, *mixture* kedua (b3[2]) -1,028, *mixture* ketiga (b3[3]) 0,1697, *mixture* keempat (b3[4]) 0,056, *mixture* kelima (b3[5]) 0,1695 dan *mixture* keenam (b3[6]) 0,2573. Pada variabel status bekerja yang menjadi acuan untuk tingkat

kecenderungan yaitu responden dengan kategori status belum bekerja, sedangkan responden kategori status sudah bekerja sebagai pembanding.

Pada *mixture* pertama, tingkat kecenderungan responden yang berstatus belum bekerja untuk memperoleh kesempatan sekolah lebih lama sebesar $\exp 3,77$ kali daripada responden yang bekerja, artinya jika ada responden dengan status bekerja pada usia sekolah dengan angka lama sekolah 3 tahun maka responden yang belum bekerja akan memiliki angka lama sekolah minimal 11 tahun. Hal tersebut dikarenakan jika seorang anak bekerja pada usia sekolah maka anak tersebut akan lebih cenderung untuk memprioritaskan pekerjaan daripada sekolah.

Pada *mixture* kedua sebesar 0,36 kali, pada *mixture* ketiga sebesar 1,18 kali, pada *mixture* keempat sebesar 1,06 kali, pada *mixture* kelima sebesar 1,18 kali, sedangkan pada *mixture* keenam sebesar 1.29 kali daripada responden yang sudah bekerja.

Kecuali pada *mixture* kedua, pada komponen *mixture* lain responden anak umur 16-18 tahun yang berstatus belum bekerja tingkat kecenderungan memiliki kesempatan sekolah lebih lama diatas 9 tahun (lulus tingkat SMP/ sederajat) lebih tinggi daripada responden yang berstatus sudah bekerja.

4. Variabel Tingkat Pendidikan Tertinggi Kepala Rumah Tangga (X4)

Variabel tingkat pendidikan kepala rumah tangga (X4) dari responden berbentuk data ordinal dengan 5 kategori, seperti yang telah dijelaskan pada bab 2, kategori 1 untuk kepala rumah tangga yang tidak tamat SD/ sederajat, kategori 2 untuk kepala rumah tangga yang tamat SD/ sederajat, kategori 3 untuk kepala rumah tangga yang tamat SMP/ sederajat, kategori 4 untuk kepala rumah tangga yang tamat SMA/ sederajat, dan kategori 5 untuk kepala rumah tangga yang tamat PT. Dalam penentuan variabel dummy untuk variabel tingkat pendidikan kepala rumah tangga, dummy 1 (X4_1) saat kategori 1 (kepala rumah tangga dengan tingkat pendidikan terakhir tidak tamat SD/ sederajat) ada dan kategori yang lain tidak ada, dummy 2 (X4_2) saat kategori 2 (kepala rumah tangga dengan tingkat pendidikan terakhir tamat SD/ sederajat) ada dan kategori yang lain tidak ada, dummy 3 (X4_3) saat kategori 3 (kepala rumah tangga dengan tingkat pendidikan terakhir tamat SMP/ sederajat) ada dan kategori yang lain tidak ada, dummy 4

(X4_4) saat kategori 4 (kepala rumah tangga dengan tingkat pendidikan terakhir tamat SMA/ sederajat) ada dan kategori yang lain tidak ada, dan saat dummy 1 hingga dummy 4 tidak ada, maka itu saat kategori 5 ada.

Variabel tingkat pendidikan kepala rumah tangga (X4) dari responden secara signifikan mempengaruhi model *mixture survival* spasial pada beberapa komponen *mixture* (*mixture* kedua, ketiga, keempat, kelima, dan keenam) dan dummy juga yang signifikan pada tiap komponen ada yang semua dummy, ada juga yang sebagian dikarenakan ada dummy yang tidak memiliki responden pada komponen *mixture* tertentu.

Pada *mixture* kedua dummy yang secara signifikan mempengaruhi dummy 2 ($b4_2[2]=0,6229$), dummy 3 ($b4_3[2]=-1,329$), dan dummy 4 ($b4_4[2]=-1,766$). Pada *mixture* ketiga dummy yang secara signifikan mempengaruhi dummy 2 ($b4_2[3]=0,2694$) dan dummy 3 ($b4_3[3]=-1,024$). Pada *mixture* keempat yang secara signifikan mempengaruhi semua dummy dengan dummy 1 ($b4_1[4]=1,294$), dummy 2 ($b4_2[4]=1,262$), dummy 3 ($b4_3[4]=1,375$), dan dummy 4 ($b4_4[4]=1,302$). Begitu juga pada *mixture* kelima dimana dummy 1 ($b4_1[5]=0,603$), dummy 2 ($b4_2[5]=0,6197$), dummy 3 ($b4_3[5]=0,67$), dan dummy 4 ($b4_4[5]=0,69$). Sedangkan pada *mixture* keenam hanya dummy 2 saja yang signifikan dengan nilai parameter $b4_2[6]=0,153$.

Pada *mixture* kedua, kepala rumah tangga dengan status pendidikan tertinggi tamat SD/ sederajat lebih cenderung memiliki anak umur 16-18 tahun (dalam hal ini responden) yang memperoleh kesempatan bersekolah lebih lama sebesar 1,86 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain. Untuk kepala rumah tangga dengan status pendidikan tertinggi tamat SMP/ sederajat lebih cenderung memiliki anak umur 16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 0,26 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain, sedangkan untuk kepala rumah tangga dengan status pendidikan tertinggi tamat SMA/ sederajat lebih cenderung memiliki anak umur 16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 0,17 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain.

Pada *mixture* ketiga, artinya kepala rumah tangga dengan status pendidikan tertinggi tamat SD/ sederajat lebih cenderung memiliki anak umur 16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 1,31 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain, sedangkan untuk kepala rumah tangga dengan status pendidikan tertinggi tamat SMP/ sederajat lebih cenderung memiliki anak umur 16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 0,36 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain.

Pada *mixture* keempat, kepala rumah tangga dengan status pendidikan tertinggi tidak tamat SD/ sederajat lebih cenderung memiliki anak umur 16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 3,65 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain, untuk kepala rumah tangga dengan status pendidikan tertinggi tamat SD/ sederajat lebih cenderung memiliki anak umur 16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 3,53 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain, untuk kepala rumah tangga dengan status pendidikan tertinggi tamat SMP/ sederajat lebih cenderung memiliki anak umur 16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 3,96 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain, sedangkan untuk kepala rumah tangga dengan status pendidikan tertinggi tamat SMA/ sederajat lebih cenderung memiliki anak umur 16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 3,68 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain.

Pada *mixture* kelima, kepala rumah tangga dengan status pendidikan tertinggi tidak tamat SD/ sederajat lebih cenderung memiliki anak umur 16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 1,83 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain, untuk kepala rumah tangga dengan status pendidikan tertinggi tamat SD/ sederajat lebih cenderung memiliki anak umur 16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 1,86 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain, untuk kepala rumah tangga dengan status pendidikan tertinggi tamat SMP/ sederajat lebih cenderung memiliki anak umur

16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 1,95 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain, sedangkan untuk kepala rumah tangga dengan status pendidikan tertinggi tamat SMA/ sederajat lebih cenderung memiliki anak umur 16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 1,95 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain.

Pada *mixture* keenam, variabel pendidikan tertinggi kepala rumah tangga dari responden untuk kepala rumah tangga dengan status pendidikan tertinggi tamat SD/ sederajat lebih cenderung memiliki anak umur 16-18 tahun yang memperoleh kesempatan bersekolah lebih lama sebesar 1,17 kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain

5. Variabel Jenis Kelamin Kepala Rumah Tangga (X5)

Variabel jenis kelamin kepala rumah tangga (X5) secara signifikan mempengaruhi model *mixture survival* spasial pada semua komponen *mixture*. Dengan nilai parameter dari variabel jenis kelamin kepala rumah tangga responden pada *mixture* pertama (b5[1]) 0,4103, *mixture* kedua (b5[2]) -0,8042, *mixture* ketiga (b5[3]) -0,2245, *mixture* keempat (b5[4]) 0,1496, *mixture* kelima (b5[5]) 0,2619 dan *mixture* keenam (b5[6]) 0,1959. Kategori yang menjadi acuan sama seperti variabel X1 yaitu kepala rumah tangga berjenis kelamin laki-laki, sedangkan kepala rumah tangga berjenis kelamin perempuan sebagai pembanding.

Pada *mixture* pertama, tingkat kecenderungan kepala rumah tangga dengan jenis kelamin laki-laki memiliki anak umur 16-18 tahun untuk memperoleh kesempatan bersekolah lebih lama sebesar 1,51 kali daripada responden dengan kepala rumah tangga berjenis kelamin perempuan, artinya jika ada responden dengan kepala rumah tangga perempuan memiliki angka lama sekolah 7 tahun, maka responden lainnya dengan kepala rumah tangga laki-laki akan memiliki angka lama sekolah minimal sebesar 8 tahun. Responden dengan kepala rumah tangga perempuan akan lebih cenderung untuk bekerja juga membantu keuangan rumah tangga sehingga untuk keperluan sekolah dikesampingkan.

Pada *mixture* kedua sebesar 0,45 kali, pada *mixture* ketiga sebesar 0,8 kali, pada *mixture* keempat sebesar 1,16 kali, pada *mixture* kelima sebesar 1,3 kali,

sedangkan pada *mixture* keenam sebesar 1,22 kali daripada responden dengan kepala rumah tangga berjenis kelamin perempuan.

6. Variabel Pengeluaran Per Kapita Per Bulan Rumah Tangga (X6)

Variabel pengeluaran per kapita per bulan rumah tangga (X6) secara signifikan mempengaruhi model *mixture survival* spasial pada *mixture* kedua ($b6[2]=0,1589$), *mixture* ketiga ($b6[3]=0,5691$), *mixture* keempat ($b6[4]=0,07035$), dan *mixture* kelima ($b6[5]=0,01332$). Seperti yang telah dijelaskan pada bab 2 bahwa interpretasi untuk variabel kontinu (data rasio dan interval) berbeda dengan interpretasi untuk variabel kategorik.

Pada *mixture* kedua, setiap penambahan 1 satuan pengeluaran per kapita per bulan rumah tangga lebih cenderung seorang responden memperoleh kesempatan bersekolah lebih lama sebesar 1,17 kali atau dalam interpretasi lain, setiap penambahan 1 satuan pengeluaran per kapita per bulan rumah tangga lebih cenderung membuat keinginan anak umur 16-18 tahun pada rumah tangga tersebut sebesar 1,17 kali untuk tetap bersekolah.

Untuk *mixture* ketiga sebesar 1,77 kali, pada *mixture* keempat sebesar 1,07 kali, sedangkan pada *mixture* kelima sebesar 1,01 kali

7. Variabel Jumlah Anggota Rumah Tangga (X7)

Variabel jumlah anggota rumah tangga (X7) secara signifikan mempengaruhi model *mixture survival* spasial pada *mixture* kedua ($b7[2]=0,7912$), *mixture* ketiga ($b7[3]=0,3746$), *mixture* keempat ($b7[4]=0,0232$), *mixture* kelima ($b7[5]=0,07189$), dan *mixture* keenam ($b7[6]=0,054$).

Pada *mixture* kedua, setiap penambahan 1 satuan anggota rumah tangga lebih cenderung membuat keinginan anak umur 16-18 tahun pada rumah tangga tersebut sebesar 2.21 kali untuk mendapatkan kesempatan bersekolah lebih sedikit atau tidak melanjutkan pendidikan untuk jenjang pendidikan yang lebih tinggi dari SMP/ sederajat.

Pada *mixture* ketiga sebesar 1.45 kali, pada *mixture* keempat sebesar 1.02 kali, pada *mixture* kelima sebesar 1.07 kali, sedangkan pada *mixture* keenam sebesar 1.06 kali.

8. Variabel Lokasi Tempat Tinggal Responden (X8)

Variabel lokasi tempat tinggal responden (X8) yaitu perkotaan atau perdesaan secara signifikan mempengaruhi model *mixture survival* spasial pada semua komponen *mixture*. Dengan nilai parameter dari variabel lokasi tempat tinggal responden pada *mixture* pertama (b8[1]) -0,8775, *mixture* kedua (b8[2]) 0,5496, *mixture* ketiga (b8[3]) 0,02, *mixture* keempat (b8[4]) 0,07, *mixture* kelima (b8[5]) 0,2039 dan *mixture* keenam (b8[6]) 0,5175. Kategori acuan yaitu responden yang berlokasi tempat tinggal di perkotaan, sedangkan responden yang berlokasi tempat tinggal di perdesaan menjadi pembanding.

Pada *mixture* pertama, tingkat kecenderungan responden dengan lokasi tempat tinggal di perkotaan untuk memperoleh kesempatan sekolah lebih lama sebesar 0,42 kali daripada responden dengan lokasi tempat tinggal di perdesaan. Pada *mixture* kedua sebesar 1,73 kali, pada *mixture* ketiga sebesar 1,02 kali, pada *mixture* keempat sebesar 1,07 kali, pada *mixture* kelima sebesar 1,23 kali, sedangkan pada *mixture* keenam sebesar 1,68 kali daripada responden dengan lokasi tempat tinggal di perdesaan.

Selain variabel-variabel yang telah disebutkan di atas, dapat diketahui bahwa parameter *frailty* CAR ($\sigma(\lambda) = 0.3562$) juga signifikan berpengaruh terhadap laju kelulusan responden anak umur 16-18 tahun, dapat dilihat. Hal ini berarti bahwa dalam kasus angka lama sekolah ini memang terdapat efek spasial dalam model *mixture survival* standar sehingga mengakibatkan munculnya efek spasial.

Tabel 4.7. Efek Spasial dari Setiap Kabupaten di Jawa Timur

Node	Mean	Sd	Mc Error	2.5%	Median	97.5%
W[1]	-0.2717	0.04546	0.006049	-0.3782	-0.2672	-0.1914
W[2]	-0.3483	0.03787	0.005593	-0.4128	-0.3565	-0.285
W[3]	0.00197	0.03312	0.002512	-0.05956	0.001607	0.06699
W[4]	-0.07598	0.02272	0.002929	-0.1185	-0.07568	-0.02693
W[5]	-0.02175	0.03081	0.004051	-0.07908	-0.0228	0.04001
W[6]	-0.009028	0.02444	0.002798	-0.05653	-0.008721	0.0343
W[7]	-0.02994	0.02822	0.003328	-0.07783	-0.03216	0.02657
W[8]	0.05817	0.03094	0.003085	9,87E-01	0.05759	0.1213
W[9]	0.0655	0.02438	0.003243	0.02274	0.065	0.1122
W[10]	-0.01428	0.03373	0.004913	-0.06975	-0.02132	0.06326
W[11]	0.04552	0.03446	0.003978	-0.0161	0.04428	0.1211
W[12]	0.0855	0.03918	0.004739	0.01705	0.08556	0.1652
W[13]	0.00452	0.02288	0.002952	-0.04237	0.004623	0.05193
W[14]	-0.05524	0.03654	0.005401	-0.1101	-0.06866	0.009866
W[15]	0.08199	0.0243	0.001635	0.0331	0.08204	0.1291
W[16]	-0.009999	0.02181	0.002487	-0.05154	-0.01001	0.03211
W[17]	0.0905	0.02432	0.0021	0.04211	0.08982	0.1366
W[18]	-0.0434	0.02463	0.003543	-0.08477	-0.04538	-0.003292
W[19]	0.01097	0.02669	0.001846	-0.04092	0.007773	0.07024
W[20]	0.04101	0.0381	0.002447	-0.03863	0.04283	0.1117
W[21]	0.03902	0.02956	0.00267	-0.01865	0.03782	0.096
W[22]	-0.2173	0.0168	0.002384	-0.249	-0.2142	-0.1918
W[23]	-0.01373	0.0255	0.002468	-0.06406	-0.01291	0.03949
W[24]	-0.1714	0.02835	0.004137	-0.2106	-0.1778	-0.1147
W[25]	0.04598	0.02255	0.001923	0.001085	0.04582	0.09086
W[26]	-0.08357	0.03369	0.00477	-0.1602	-0.08514	-0.023
W[27]	-0.1708	0.02792	0.003044	-0.2244	-0.1722	-0.1075
W[28]	-0.06221	0.02883	0.002299	-0.1202	-0.06231	-0.006239
W[29]	0.03104	0.03067	0.001958	-0.02994	0.03154	0.09379
W[30]	0.137	0.03376	0.00326	0.07085	0.1384	0.2001
W[31]	0.1149	0.03634	0.001972	0.04691	0.1154	0.1807
W[32]	0.1578	0.02699	0.003622	0.1054	0.1563	0.2083
W[33]	0.1685	0.0288	0.001633	0.1138	0.1675	0.2249
W[34]	0.1268	0.02978	0.00356	0.07493	0.1266	0.1816
W[35]	0.1157	0.03315	0.003548	0.05803	0.1146	0.1899
W[36]	0.12	0.0342	0.002553	0.04685	0.1217	0.1847
W[37]	0.07523	0.02374	0.00184	0.03056	0.07608	0.1176
W[38]	-0.01891	0.03777	0.003184	-0.09832	-0.02053	0.05435

Sumber : Perhitungan WinBUGS

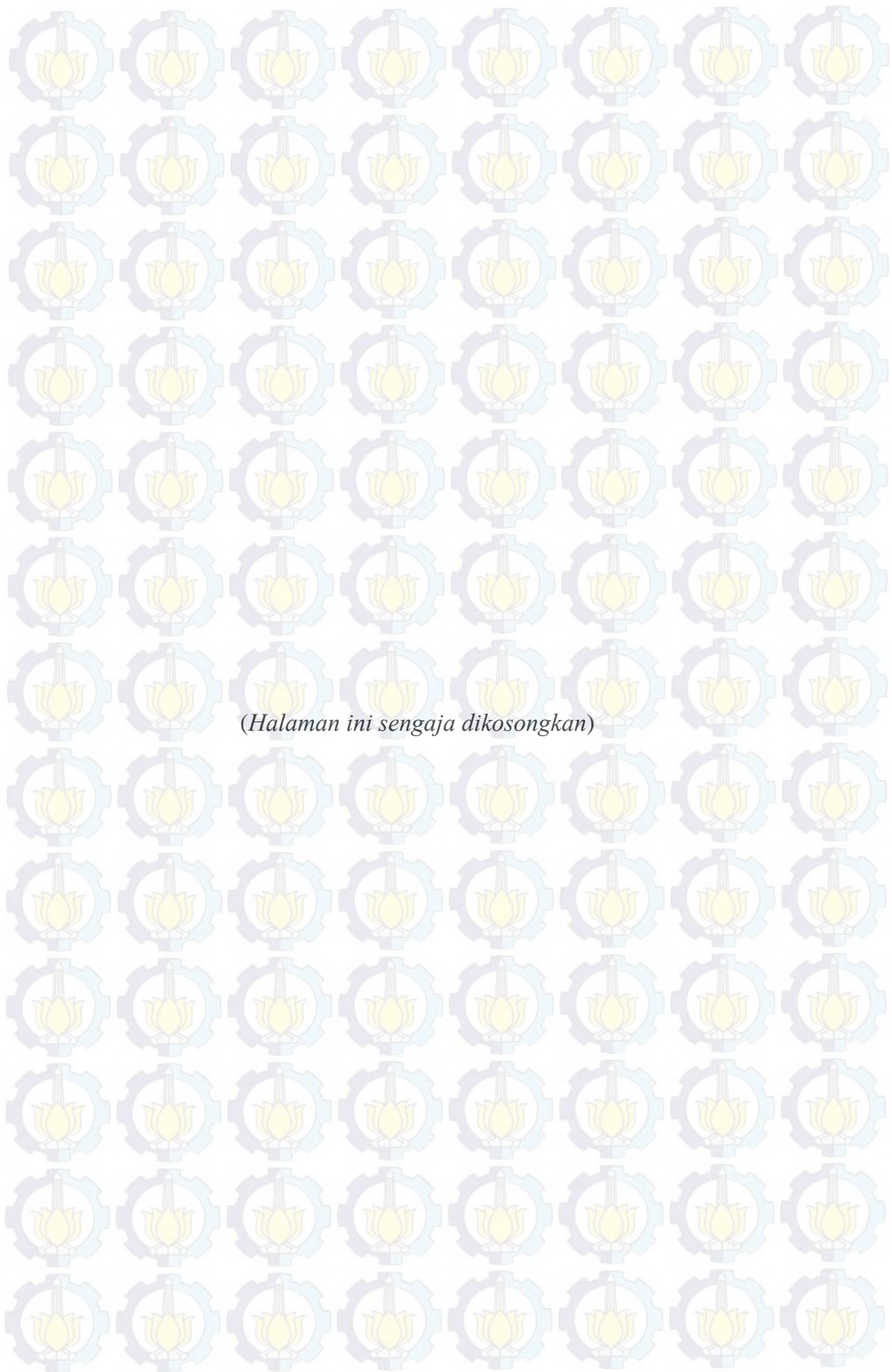
Pada Tabel 4.7. dapat dilihat bahwa baris yang berwarna kuning adalah daerah-daerah dengan efek spasial yang tidak signifikan, daerah-daerah tersebut adalah Trenggalek, Blitar, Kediri, Malang, Banyuwangi, Bondowoso, Probolinggo, Pasuruan, Mojokerto, Madiun, Magetan, Ngawi, Tuban, Sumenep, dan Kota Batu. sedangkan yang berwarna putih adalah daerah-daerah efek spasial yang signifikan, daerah-daerah tersebut yaitu Pacitan, Ponorogo, Tulungagung, Lumajang, Jember, Situbondo, Sidoarjo, Jombang, Nganjuk, Bojonegoro, Lamongan, Gresik, Bangkalan, Sampang, Pamekasan, Kota Kediri, Kota Blitar, Kota Malang, Kota Probolinggo, Kota Pasuruan, Kota Mojokerto, Kota Madiun, dan Kota Surabaya. Signifikan atau tidaknya efek spasial dapat dilihat dari nilai pada selang 2.5% hingga 97.5% tidak memuat nilai 0.

Daerah-daerah dengan efek spasial yang secara signifikan mempengaruhi angka lama sekolah anak umur 16-18 tahun artinya bahwa pengaruh spasial atau pengaruh kedaerahan, dimana dalam penelitian ini pengaruh spasial dianggap random (menjadi variabel random W_j^*), memiliki efek yang berbeda di setiap daerah dan setiap daerah mungkin saja memiliki kebijakan yang berbeda atau keadaan yang berbeda yang diciptakan oleh masyarakat antara satu daerah dengan daerah lain sehingga mempengaruhi angka lama sekolah anak umur 16-18 tahun di daerah tersebut yang berimplikasi bahwa rata-rata lama sekolah berdasarkan angka lama sekolah anak umur 16-18 tahun di daerah tersebut berbeda juga.

Berdasarkan Tabel 4.5 dan Tabel 4.7, maka model mixture survival spasial yang terbentuk yaitu

$$\begin{aligned}
h(t, \mathbf{x}) = & 0,0087(h_1(t) \exp(0,98x_2 + 1,33x_3 + 0,41x_5 + 0,88x_8 + W_j^*)) \\
& + 0,008(h_2(t) \exp(0,53x_1 + 0,38x_2 - 1,03x_3 + 0,62x_{4_2} \\
& - 1,33x_{4_3} - 1,77x_{4_4} - 0,8x_5 + 0,16x_6 + 0,8x_7 + 0,55x_8 + W_j^*)) \\
& + 0,06(h_3(t) \exp(0,54x_1 + 0,25x_2 + 0,17x_3 + 0,27x_{4_2} - 1,02x_{4_3} \\
& - 0,22x_5 + 0,57x_6 + 0,37x_7 + 0,02x_8 + W_j^*)) \\
& + 0,092(h_4(t) \exp(0,1x_1 + 0,34x_2 + 0,06x_3 + 1,29x_{4_1} + 1,26x_{4_2} \\
& + 1,38x_{4_3} + 1,3x_{4_4} + 0,15x_5 + 0,07x_6 + 0,02x_7 + 0,07x_8 + W_j^*)) \\
& + 0,798(h_5(t) \exp(0,14x_1 + 0,23x_2 + 0,17x_3 + 0,6x_{4_1} + 0,62x_{4_2} \\
& + 0,67x_{4_3} + 0,69x_{4_4} + 0,26x_5 + 0,013x_6 + 0,072x_7 + 0,2x_8 + W_j^*)) \\
& + 0,029(h_6(t) \exp(0,122x_1 + 0,73x_2 + 0,26x_3 + 0,15x_{4_2} + 0,2x_5 \\
& + 0,05x_7 + 0,52x_8 + W_j^*))
\end{aligned}$$

dimana $h_1(t)$ hingga $h_6(t)$ adalah fungsi baseline hazard dari setiap komponen mixture dan W_j^* adalah efek spasial dari daerah ke- j . Nilai $h(t, \mathbf{x})$ adalah nilai besarnya laju atau kesempatan untuk bersekolah lebih lama dengan dipengaruhi oleh besar komponen, nilai baseline hazard, variabel-variabel yang secara signifikan mempengaruhi pada komponen tertentu dan daerah yang memiliki efek spasial yang secara signifikan mempengaruhi.



BAB 5

KESIMPULAN DAN SARAN

Pada bab ini disimpulkan mengenai hasil dan pembahasan yang telah dijelaskan pada bab sebelumnya dan untuk menjawab rumusan masalah pada bab pertama serat memberikan saran yang membangun untuk manfaatnya penelitian ini maupun sebagai acuan untuk penelitian selanjutnya.

5.1. Kesimpulan

Berdasarkan hasil dan pembahasan pada bab sebelumnya, maka dapat disimpulkan,

1. Model umum *mixture survival* spasial dimana pendekatan setiap komponen *mixture* dengan distribusi normal didapatkan 6 komponen *mixture* dan efek spasial W_j^* didapatkan sebagai berikut,

$$\begin{aligned} h(t, \mathbf{x}) = & \pi_1(h_1(t) \exp(\beta_1^T x_{p1} + W_j^*)) + \pi_2(h_2(t) \exp(\beta_2^T x_{p2} + W_j^*)) \\ & + \pi_3(h_3(t) \exp(\beta_3^T x_{p3} + W_j^*)) + \pi_4(h_4(t) \exp(\beta_4^T x_{p4} + W_j^*)) \\ & + \pi_5(h_5(t) \exp(\beta_5^T x_{p5} + W_j^*)) + \pi_6(h_6(t) \exp(\beta_6^T x_{p6} + W_j^*)) \end{aligned}$$

dimana,

π_1, \dots, π_6 adalah komponen *mixture* pertama hingga keenam

$h_1(t), \dots, h_6(t)$ adalah fungsi baseline *hazard* pertama hingga keenam dimana memuat t dan parameter σ distribusi normal.

$\beta^T x_p$ adalah komponen linier dari model yang memuat variabel-variabel yang signifikan disetiap komponen *mixture*
 W_j^* adalah efek random spasial kabupaten/kota ke- j

2. Variabel-variabel yang berpengaruh secara signifikan berbeda pada tiap komponen *mixture*
 - a. Komponen *mixture* pertama, besar proporsi 0.00869, responden yang masuk dengan angka lama sekolah 0, 1 dan 2. Variabel-variabel yang mempengaruhi model secara signifikan meliputi status perkawinan

responden (X2), status bekerja responden (X3), jenis kelamin kepala rumah tangga (X5), dan lokasi tempat tinggal responden (X8).

b. Komponen *mixture* kedua, besar proporsi 0.00767, responden yang masuk angka lama sekolah 1, 2, 3, 4, dan 5. Variabel-variabel yang mempengaruhi model secara signifikan meliputi semua variabel kecuali variabel dummy kategori 1 pada variabel tingkat pendidikan tertinggi kepala keluarga.

c. Komponen *mixture* ketiga, besar proporsi 0.0642, responden yang masuk dengan angka lama sekolah 4, 5, 6, 7, dan 8. Variabel-variabel yang mempengaruhi secara signifikan meliputi semua variabel kecuali variabel dummy kategori 1 dan kategori 4 pada variabel tingkat pendidikan tertinggi kepala keluarga.

d. Komponen *mixture* keempat, besar proporsi 0.09174, responden yang masuk dengan angka lama sekolah 6, 7, 8, 9, dan 10. Semua variabel secara signifikan mempengaruhi model pada *mixture* keempat.

e. Komponen *mixture* kelima, besar proporsi 0.7984, responden yang masuk dengan angka lama sekolah 8, 9, 10, 11, dan 12, dimana sama halnya seperti *mixture* keempat dimana semua variabel secara signifikan mempengaruhi model pada *mixture* kelima.

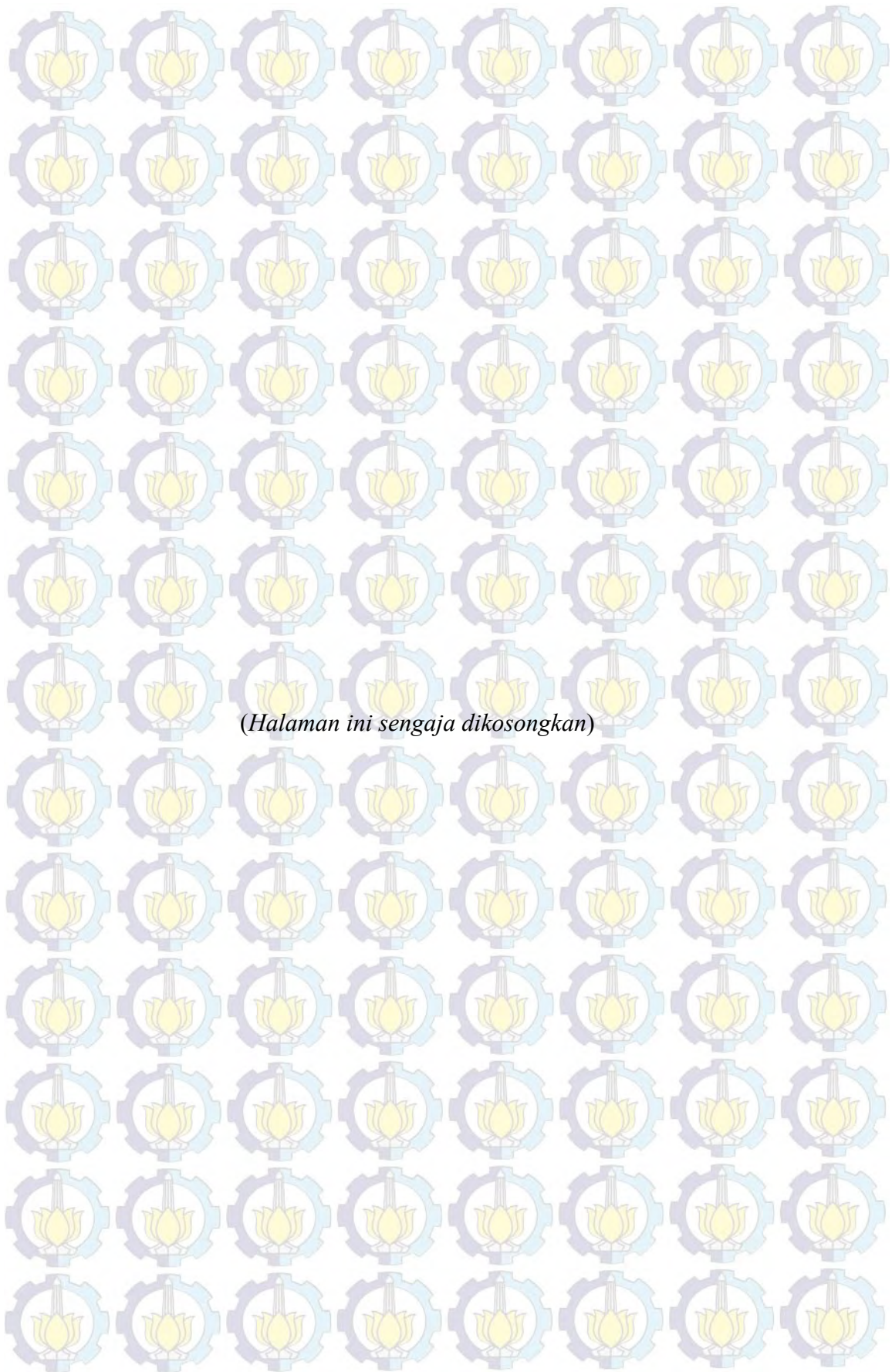
f. Sedangkan untuk *mixture* keenam dengan besar proporsi 0.02929, responden yang masuk dengan angka lama sekolah 10, 11, dan 12. Variabel-variabel yang mempengaruhi model secara signifikan meliputi jenis kelamin responden (X1), status perkawinan responden (X2), status bekerja responden (X3), jenis kelamin kepala rumah tangga (X5), jumlah anggota rumah tangga (X7), dan lokasi tempat tinggal responden (X8).

g. Serta untuk efek spasial yang mempengaruhi model secara signifikan yaitu Pacitan, Ponorogo, Tulungagung, Lumajang, Jember, Situbondo, Sidoarjo, Jombang, Nganjuk, Bojonegoro, Lamongan, Gresik, Bangkalan, Sampang, Pamekasan, Kota Kediri, Kota Blitar, Kota Malang, Kota Probolinggo, Kota Pasuruan, Kota Mojokerto, Kota Madiun, dan Kota Surabaya.

5.2. Saran

Adapun saran yang diajukan dalam penelitian ini adalah

1. Pada penelitian selanjutnya diharapkan agar data untuk model mixture survival spasial menggunakan data panel agar selain ada efek spasial terdapat juga efek waktu yang mempengaruhi variabel respon.
2. Pada penelitian selanjutnya perlu pengembangan dalam bidang spasial, agar yang digunakan tidak hanya efek spasial namun dependensi spasial.
3. Perlu juga dilakukan pengembangan dalam Reversible Jump Markov Chain Monte Carlo untuk distribusi selain distribusi Normal.



Daftar Pustaka

- Aksioma, D. F. (2011). *Model Spasial Survival dengan Pendekatan Bayesian (Studi Kasus pada Kejadian HIV/AIDS di Provinsi Jawa Timur)*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Aksioma, D. F., & Iriawan, N. (2010). Spatial Autocorrelation of the DHF Outbreaks in the City of Surabaya. *Proceeding of The Third International Conference on Mathematics and Natural Science (ICMNS) 2010*. Bandung.
- Ando, T., Imoto, S., & Miyano, S. (2004). Kernel Mixture Survival Models for Identifying Cancer Subtypes, Predicting Patient's Cancer Types and Survival Probabilities. *Genome Informatics*, 15 (2), 201-210.
- Banerjee, S., Wall, M. M., & Carlin, B. P. (2003). Frailty Modeling for Spatially Correlated Survival Data, With Application to Infant Mortality in Minnesota. *Biostatistics*, 123-142.
- Besag, J., York, J., & Mollie, A. (1991). Bayesian Image Restoration with Two Application in Spatial Statistics. *Annals of the Institute of Statistical Mathematics*, 1-59.
- Best, N., Mason, A., dan Li, P. (2011), *Bayesian hierarchical modelling using WinBUGS*. Economic and Social Research Council, ESRC National Centre for Research Methods, UK.
- Boldstad, W.M., (2007). *Introduction to Bayesian Statistics. 2nd Edition*. New Jersey: Wiley.
- Box, G. E. & Tiao, G. C., (1973). *Bayesian Inference in Statistical Alaysis*. London: Addison-Wisley.
- BPS. (2014). *sirusa.bps.go.id*. Retrieved October 6, 2014, from sirusa.bps.go.id: <http://sirusa.bps.go.id>
- BPS. (2012). *Statistik Indonesia 2012*. Jakarta: BPS.
- BPS. (2012). *Pedoman Survei Sosial Ekonomi Nasional 2013*. Jakarta: BPS.
- Carlin, B. P., & Louis, T. A. (2000). *Bayes And Empirical Bayes Methods For Data Analysis (2ed)*. Boca Raton, FL: Chapman and Hall/CRC Press.
- Casella, G. dan George, E.I. (1992), Explaining the Gibbs Sampler. *Journal of the American Statistical Association* 46(3): 167-174.
- Collet, D. (1994). *Modelling Survival Data in Medical Research*. London: Chapman and Hall.
- Collet, D. (2003). *Modelling Survival Data in Medical Research 2nd ed*. London: Chapman and Hall.
- Cox, D., & Oakes, D. (1984). *Analysis of Survival Data*. London: Chapman and Hall.

- Cressie, N. A. C. (1993). *Statistics for Spatial Data*, revised edition. New York: Wiley.
- Dalyono. (2008). *Psikologi Pendidikan*. Jakarta: Rineka Cipta.
- Darmofal, D. (2008). *Bayesian Spatial Survival Models for Political Event Processes*. Columbia: Departement of Political Science, University of South Carolina 350 Gambrel Hall.
- Ducrocq, V. (1997). *Survival Analysis, A Statistical Tool For Longevity Data, 48th Annual Meeting Of The European Assosiation For Animal Production*. Vienna: Institut National de la Recherche Agronomique.
- Gamerman, D., 1997. *Markov Chain Monte Carlo*. London: Chapman & Hall.
- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995). *Bayesian Data Analysis* (2nd ed.). London: Chapman and Hall.
- Gelman, A., & Pardoe, L. (2006), *Bayesian Measure of Explained Variance and Pooling in Multilevel (Hierarchical) Models*, *Technometrics*, **48**(2): 241-251.
- Green, P. J., 1995. Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination. *Biometrika*, 82(4), pp. 711-732.
- Hariyanto, S. (2009). *Model Mixture Survival pada Kasus Lama Mencari Kerja di Pulau Jawa Tahun 2007, Tesis*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Hasyim, M. (2012). *Model Mixture Survival Spasial dengan Frailty Berdistribusi Conditionally Autoregressive (CAR) pada Kasus Kejadian Demam Berdarah Dengue (DBD) di Kabupaten Pamekasan*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Hobcraft, J., McDonald, J., Menken, J., Rodriguez, G., & Trussel, J. (1984). *A Comparative Analysis of determinants of Birth Intervals. In, WFS Comparative Study: Cross-National Summaries (World Fertility Surveys)*. Voorburg, Netherland: International Statistical Institute.
- Hosmer Jr., D., & Lemeshow, S., .. (1999). *Applied Survival Analysis Regression Modelling of Time to Event Data*. New York: John Wiley and Sons. Inc.
- Ikana, M. (2005). *Pengaruh Urutan Kelahiran Terhadap Kelangsungan Pendidikan Anak Perempuan Usia 7-15 Tahun di Indonesia (Analisis Data Susenas 2002-KOR), Tesis*. Depok: Universitas Indonesia.
- Iriawan, N. (2000). Pemilihan Model Menggunakan Struktur Perkalian Distribusi. *KAPPA: Jurnal Sains*, 37-44.
- Iriawan, N. (2001). *Studi Tentang Bayesian Mixture Normal dengan Menggunakan Metode MCMC*. Surabaya: Lemlit ITS.
- Ismartini, P. (2013). *Pengembangan Model Linear Hirarki Dengan Pendekatan Bayesian Untuk Pemodelan Data Pengeluaran Per Kapita Rumah Tangga*. Surabaya: Institut Teknologi Sepuluh Nopember.

- Kartika, Y. (2007). *Pola Penyebaran Spasial Demam Berdarah Dengue di Kota Bogor Tahun 2005. Tugas Akhir*. Bogor: Institut Pertanian Bogor.
- Kleinbaum, D. G., & Klein, M. (2005). *Survival Analysis: A Self Learning* (3rd ed.). New York: Springer.
- Law, A. M., & Kelton, D. W. (2000). *Simulation Modeling and Analysis* (3rd ed) (3rd ed.). New York: McGraw-Hill.
- Lee, E. (1992). *Statistical Models and Methods for Lifetime Data*. New York: John Wiley and Sons. Inc.
- Lesage, J. P. (1999). *The Theory and Practice of Spatial Econometrics*. Toledo: Department of Economics, University of Toledo.
- Li, Y., & Ryan, L. (2002). Modelling Spatial Survival Data Using Semiparametric Frailty Models. *Biometrics*, 58, 287-297.
- Marin, J. M., Mengersen, K. & Robert, C. P., (2001). Bayesian Modelling and Indereence on Mixture of Distribution. *Handbook of Statistics*, 25(50).
- McLachlan, G., & Basford, K. E. (1988). *Mixture Models Inference and Application to Clustering*. New York: Marcel Dekker.
- McLachlan, G., & Peel, D. (2000). *Finite Mixture Models*. New York: John Wiley and Sons, Inc.
- Miller, R. (1998). *Survival Analysis*. New York: John Willey and Sons Inc.
- Muthen, B., & Masyn, K. (2005). Discrete-Time Mixture Survival Analysis. *Journal of Educational and Behavioral Statistics*, 30(1), 27-58.
- Ntzoufras, I. (2009). *Bayesian Modelling Using WinBUGS*. New York: John Willey and Sons, Inc.
- Richardson, S., & Green, P. J. (1997). On Bayesian Analysis of Mixture with an Unknown Number of Component. *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 59, No. 4, 731-792.
- Santoso, B. (2009). *Pendekatan Spline Multivariable Dan MARS Untuk Pemodelan Lama Sekolah Pada Penduduk Usia Sekolah di Provinsi Papua, Tesis*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P., and Linde, A. (2002), *Bayesian measures of model complexity and fit*. *Journal of The Royal Statistical Society*, 64(1): 583-639.
- Suendra, N. (1999). *Pemetaan Peserta Didik Wajib Belajar 9 Tahun yang Putus Sekolah dan Tidak Melanjutkan Ke SLTP Pada Desa-Desa Tertinggal di Provinsi Bali, Tugas Akhir*. Bali: Jurusan Ilmu Pendidikan dan Keguruan Universitas Singaraja.
- Sulistiyawati, D. (2009). *Model Mixture Survival Pada Kasus Lama Sekolah di Kabupaten Boalemo Provinsi Gorontalo, Tesis*. Surabaya: Institut Teknologi Sepuluh Nopember.

Vittinghoff, E., Glidden, D. V., Shiboski, S. C., & McCulloch, C. E. (2005). *Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models*. New York: Springer.

Wackerly, D.D., Mendenhall, W., dan Scheaffer, R.L. (1996), *Mathematical Statistics with Applications*. Duxbury Press, California.

Wall, M. M. (2004). A Close Look at The Spatial Structure Implied by the CAR and SAR Models. *Journal of Statistical Planning and Inference*, 121 , 311-324.

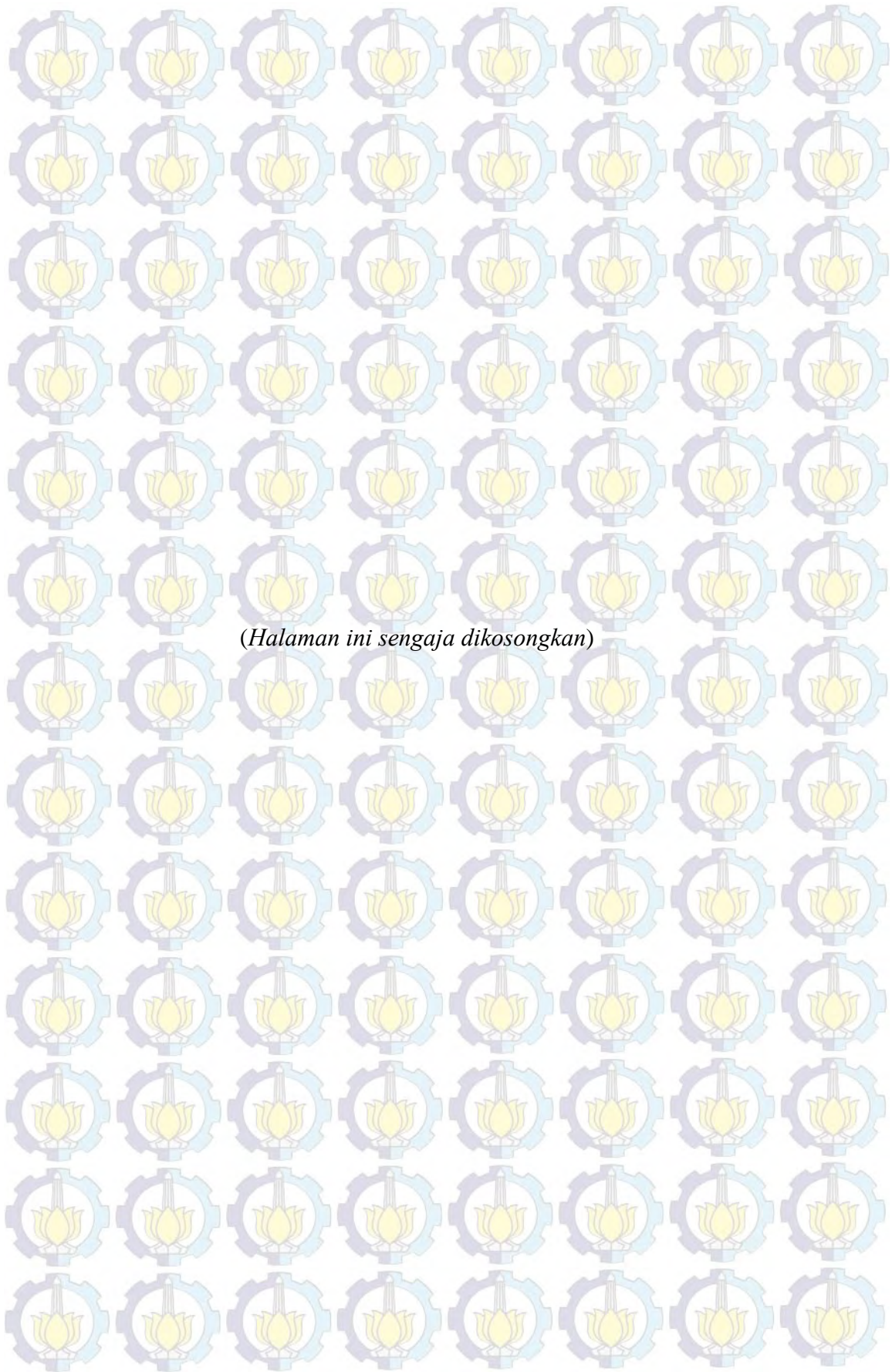
Zang. (2008). *Survival Analysis*. California: Wadsworth.

Zhang, Y. (2008). *Parametric Mixture Models in Survival Analysis with Application, Disertation*. Philadelphia: Temple University.



DAFTAR LAMPIRAN

Lampiran 1. Data Angka Anak Umur 16-18 Tahun se Jawa Timur.....	85
Lampiran 2. Hasil Pengujian Distribusi Data Angka Lama Sekolah.....	86
Lampiran 3. Hasil Generalized Model.....	87
Lampiran 4. Syntax Reversible Jump MCMC.....	93
Lampiran 5. Hasil Syntax Reversible Jump MCMC	94
Lampiran 6. Syntax Gibbs Sampling Mixture Normal 6 Komponen	96
Lampiran 7. Hasil Syntax Gibbs Sampling Mixture Normal 6 Komponen.....	98



Lampiran 1

No.	T	l.cen	X1	X2	X3	X4	X5	X6	X6	X7	Kab
	ALS	Sensor	JK	SP	SB	TPTKK	JKKK	Rata-Rata Peng- (dalam 00.000)	Jumlah ART	Lokasi	Kot
1.	10	Tidak tersensor	Perempuan	Belum kawin	Tidak	S1	Laki-laki	20,22	4	Kota	1
2.	9	Tidak tersensor	Laki-laki	Belum kawin	Tidak	SMA/Sederajat	Laki-laki	2,75	4	Kota	1
3.	6	Tersensor	Laki-laki	Belum kawin	Ya	SD/Sederajat	Laki-laki	1,82	5	Desa	1
4.	11	Tidak tersensor	Laki-laki	Belum kawin	Ya	SD/Sederajat	Laki-laki	2,00	9	Desa	1
5.	8	Tersensor	Laki-laki	Belum kawin	Tidak	SMA/Sederajat	Laki-laki	5,26	4	Desa	1

1990.	8	Tersensor	Perempuan	Belum kawin	Tidak	SD/Sederajat	Laki-laki	2,10	4	Desa	19
1991.	9	Tidak tersensor	Laki-laki	Belum kawin	Tidak	SD/Sederajat	Perempuan	5,61	3	Desa	19
1992.	9	Tidak tersensor	Perempuan	Belum kawin	Tidak	SD/Sederajat	Laki-laki	2,40	4	Desa	19
1993.	8	Tersensor	Perempuan	Belum kawin	Tidak	SMA/Sederajat	Laki-laki	8,17	4	Desa	19
1994.	11	Tidak tersensor	Perempuan	Belum kawin	Tidak	SMA/Sederajat	Laki-laki	4,10	5	Desa	19
1995.	9	Tidak tersensor	Laki-laki	Belum kawin	Tidak	Tdk punya ijazah SD	Laki-laki	3,62	4	Kota	20
1996.	1	Tersensor	Perempuan	Belum kawin	Ya	Tdk punya ijazah SD	Laki-laki	2,43	4	Kota	20

3790.	9	Tidak tersensor	Perempuan	Kawin	Ya	SMA/Sederajat	Laki-laki	4,19	5	Desa	38
3791.	9	Tidak tersensor	Perempuan	Belum kawin	Ya	SD/Sederajat	Laki-laki	2,77	4	Desa	38
3792.	10	Tidak tersensor	Perempuan	Belum kawin	Tidak	SMA/Sederajat	Laki-laki	6,40	5	Desa	38
3793.	9	Tidak tersensor	Perempuan	Kawin	Tidak	Tdk punya ijazah SD	Laki-laki	9,90	5	Desa	38
3794.	9	Tidak tersensor	Perempuan	Kawin	Tidak	SD/Sederajat	Laki-laki	6,29	5	Desa	38

Lampiran 2

Goodness of Fit - Summary

#	<u>Distribution</u>	<u>Kolmogorov Smirnov</u>		<u>Anderson Darling</u>	
		Statistic	Rank	Statistic	Rank
1	<u>Exponential</u>	0,45232	8	1235,0	9
2	<u>Exponential (2P)</u>	0,45232	9	1186,7	8
3	<u>Log-Logistic</u>	0,30631	7	523,32	7
4	<u>Log-Logistic</u>	0,17449	2	193,5	3
5	<u>Lognormal</u>	0,26474	4	281,78	4
6	<u>Lognormal</u>	0,26476	5	281,79	5
7	<u>Normal</u>	0,22618	3	146,06	1
8	<u>Weibull</u>	0,29574	6	517,59	6
9	<u>Weibull</u>	0,16724	1	147,0	2

Lampiran 3

Hasil Generalized Model Untuk Komponen Mixture Pertama Dengan Enam Komponen

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	1.433	1.8526	-2.198	5.064	.599	1	.439
[Jenis_Kelamin=1]	-.167	.4649	-1.078	.744	.129	1	.719
[Jenis_Kelamin=2]	0 ^a
[Status_Perkawinan=1]	.548	.8928	-1.202	2.298	.377	1	.539
[Status_Perkawinan=2]	0 ^a
[Status_Bekerja=1]	-.625	.4604	-1.527	.277	1.843	1	.175
[Status_Bekerja=2]	0 ^a
[TPKRT=1]	-.284	1.0205	-2.285	1.716	.078	1	.781
[TPKRT=2]	.098	1.0407	-1.942	2.138	.009	1	.925
[TPKRT=3]	-.188	1.0250	-2.197	1.821	.033	1	.855
[TPKRT=4]	0 ^a
[JKKRT=1]	-.111	.8911	-1.857	1.636	.015	1	.901
[JKKRT=2]	0 ^a
[Lokasi_Tempat_Tinggal=1]	-.296	.4421	-1.162	.571	.447	1	.504
[Lokasi_Tempat_Tinggal=2]	0 ^a
Rata_Peng	-.065	.1502	-.360	.229	.189	1	.664
Jumlah_ART	-.087	.1165	-.315	.142	.554	1	.457
(Scale)	1 ^b						

Dependent Variable: Angka_Lama_Sekolah

Model: (Intercept), Jenis_Kelamin, Status_Perkawinan, Status_Bekerja, TPKRT, JKKRT, Lokasi_Tempat_Tinggal, Rata_Peng, Jumlah_ART

a. Set to zero because this parameter is redundant.

b. Fixed at the displayed value.

Hasil Generalized Model Untuk Komponen Mixture Kedua Dengan Enam Komponen

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	3.484	2.3849	-1.191	8.158	2.134	1	.144
[Jenis_Kelamin=1]	.431	.5605	-.668	1.530	.591	1	.442
[Jenis_Kelamin=2]	0 ^a
[Status_Perkawinan=1]	-.474	.7803	-2.004	1.055	.370	1	.543
[Status_Perkawinan=2]	0 ^a
[Status_Bekerja=1]	-.182	.5459	-1.252	.888	.111	1	.739
[Status_Bekerja=2]	0 ^a
[TPKRT=1]	-.195	1.4142	-2.967	2.576	.019	1	.890
[TPKRT=2]	.789	1.3027	-1.764	3.342	.367	1	.545
[TPKRT=3]	-1.363	2.0335	-5.349	2.622	.449	1	.503
[TPKRT=4]	-1.755	1.7678	-5.220	1.710	.985	1	.321
[TPKRT=5]	0 ^a
[JKKRT=1]	-.551	.6868	-1.897	.795	.644	1	.422
[JKKRT=2]	0 ^a
[Lokasi_Tempat_Tinggal=1]	.182	.5016	-.801	1.165	.132	1	.717
[Lokasi_Tempat_Tinggal=2]	0 ^a
Rata_Peng	.120	.1431	-.160	.401	.709	1	.400
Jumlah_ART	-.056	.1465	-.343	.231	.146	1	.703
(Scale)	1 ^b						

Dependent Variable: Angka_Lama_Sekolah

Model: (Intercept), Jenis_Kelamin, Status_Perkawinan, Status_Bekerja, TPKRT, JKKRT, Lokasi_Tempat_Tinggal, Rata_Peng, Jumlah_ART

a. Set to zero because this parameter is redundant.

b. Fixed at the displayed value.

Hasil Generalized Model Untuk Komponen Mixture Ketiga Dengan Enam Komponen

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	6.021	.7917	4.469	7.572	57.825	1	.000
[Jenis_Kelamin=1]	.046	.1424	-.233	.326	.106	1	.744
[Jenis_Kelamin=2]	0 ^a
[Status_Perkawinan=1]	-.015	.2139	-.435	.404	.005	1	.943
[Status_Perkawinan=2]	0 ^a
[Status_Bekerja=1]	.228	.1372	-.041	.497	2.763	1	.096
[Status_Bekerja=2]	0 ^a
[TPKRT=1]	-.091	.6328	-1.331	1.149	.021	1	.886
[TPKRT=2]	.208	.6289	-1.025	1.440	.109	1	.741
[TPKRT=3]	.379	.7561	-1.103	1.861	.251	1	.617
[TPKRT=4]	.370	.7573	-1.114	1.855	.239	1	.625
[TPKRT=5]	0 ^a
[JKKRT=1]	-.050	.2102	-.462	.362	.057	1	.811
[JKKRT=2]	0 ^a
[Lokasi_Tempat_Tinggal=1]	-.195	.1516	-.492	.102	1.650	1	.199
[Lokasi_Tempat_Tinggal=2]	0 ^a
Rata_Peng	-.011	.0477	-.105	.082	.057	1	.812
Jumlah_ART	-.015	.0494	-.111	.082	.086	1	.769
(Scale)	1 ^b						

Dependent Variable: Angka_Lama_Sekolah

Model: (Intercept), Jenis_Kelamin, Status_Perkawinan, Status_Bekerja, TPKRT, JKKRT, Lokasi_Tempat_Tinggal, Rata_Peng, Jumlah_ART

a. Set to zero because this parameter is redundant.

b. Fixed at the displayed value.

Hasil Generalized Model Untuk Komponen Mixture Keempat Dengan Enam Komponen

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	8.094	.5908	6.936	9.252	187.678	1	.000
[Jenis_Kelamin=1]	-.174	.1152	-.400	.051	2.291	1	.130
[Jenis_Kelamin=2]	0 ^a
[Status_Perkawinan=1]	1.018	.2059	.615	1.422	24.444	1	.000
[Status_Perkawinan=2]	0 ^a
[Status_Bekerja=1]	.005	.1239	-.238	.248	.002	1	.965
[Status_Bekerja=2]	0 ^a
[TPKRT=1]	-.360	.4594	-1.260	.541	.613	1	.434
[TPKRT=2]	-.324	.4595	-1.225	.576	.498	1	.480
[TPKRT=3]	-.387	.4714	-1.311	.537	.673	1	.412
[TPKRT=4]	-.377	.4795	-1.316	.563	.617	1	.432
[TPKRT=5]	0 ^a
[JKKRT=1]	-.040	.2038	-.440	.359	.039	1	.843
[JKKRT=2]	0 ^a
[Lokasi_Tempat_Tinggal=1]	.032	.1180	-.199	.263	.075	1	.785
[Lokasi_Tempat_Tinggal=2]	0 ^a
Rata_Peng	.030	.0162	-.002	.062	3.413	1	.065
Jumlah_ART	-.216	.0367	-.288	-.144	34.619	1	.000
(Scale)	1 ^b						

Dependent Variable: Angka_Lama_Sekolah

Model: (Intercept), Jenis_Kelamin, Status_Perkawinan, Status_Bekerja, TPKRT, JKKRT, Lokasi_Tempat_Tinggal, Rata_Peng, Jumlah_ART

a. Set to zero because this parameter is redundant.

b. Fixed at the displayed value.

Hasil Generalized Model Untuk Komponen Mixture Kelima Dengan Enam Komponen

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	9.791	.1611	9.475	10.106	3692.173	1	.000
[Jenis_Kelamin=1]	-.167	.0369	-.240	-.095	20.485	1	.000
[Jenis_Kelamin=2]	0 ^a
[Status_Perkawinan=1]	.605	.1019	.405	.804	35.217	1	.000
[Status_Perkawinan=2]	0 ^a
[Status_Bekerja=1]	-.106	.0507	-.206	-.007	4.399	1	.036
[Status_Bekerja=2]	0 ^a
[TPKRT=1]	-.241	.0807	-.399	-.082	8.894	1	.003
[TPKRT=2]	-.117	.0755	-.265	.031	2.414	1	.120
[TPKRT=3]	.112	.0825	-.049	.274	1.852	1	.174
[TPKRT=4]	.028	.0755	-.120	.176	.139	1	.709
[TPKRT=5]	0 ^a
[JKKRT=1]	.056	.0620	-.066	.177	.813	1	.367
[JKKRT=2]	0 ^a
[Lokasi_Tempat_Tinggal=1]	.056	.0401	-.023	.134	1.936	1	.164
[Lokasi_Tempat_Tinggal=2]	0 ^a
Rata_Peng	.004	.0040	-.003	.012	1.208	1	.272
Jumlah_ART	-.064	.0145	-.092	-.035	19.363	1	.000
(Scale)	1 ^b						

Dependent Variable: Angka_Lama_Sekolah

Model: (Intercept), Jenis_Kelamin, Status_Perkawinan, Status_Bekerja, TPKRT, JKKRT, Lokasi_Tempat_Tinggal, Rata_Peng, Jumlah_ART

a. Set to zero because this parameter is redundant.

b. Fixed at the displayed value.

Hasil Generalized Model Untuk Komponen Mixture Keenam Dengan Enam Komponen

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	12.569	1.5444	9.542	15.596	66.236	1	.000
[Jenis_Kelamin=1]	.048	.2036	-.351	.447	.055	1	.814
[Jenis_Kelamin=2]	0 ^a
[Status_Perkawinan=1]	-1.171	1.0721	-3.272	.930	1.193	1	.275
[Status_Perkawinan=2]	0 ^a
[Status_Bekerja=1]	-.987	.3027	-1.581	-.394	10.643	1	.001
[Status_Bekerja=2]	0 ^a
[TPKRT=1]	-.034	.4601	-.936	.868	.006	1	.941
[TPKRT=2]	.005	.3856	-.751	.761	.000	1	.990
[TPKRT=3]	.154	.4274	-.684	.992	.130	1	.719
[TPKRT=4]	-.029	.3484	-.711	.654	.007	1	.935
[TPKRT=5]	0 ^a
[JKKRT=1]	-.113	.3450	-.789	.564	.106	1	.744
[JKKRT=2]	0 ^a
[Lokasi_Tempat_Tinggal=1]	.366	.7329	-1.070	1.803	.249	1	.617
[Lokasi_Tempat_Tinggal=2]	0 ^a
Rata_Peng	.005	.0187	-.032	.042	.072	1	.789
Jumlah_ART	.039	.0803	-.118	.197	.240	1	.624
(Scale)	1 ^b						

Dependent Variable: Angka_Lama_Sekolah

Model: (Intercept), Jenis_Kelamin, Status_Perkawinan, Status_Bekerja, TPKRT, JKKRT, Lokasi_Tempat_Tinggal, Rata_Peng, Jumlah_ART

a. Set to zero because this parameter is redundant.

b. Fixed at the displayed value.

Lampiran 4

Syntax Reversible Jump MCMC

```
model {  
  for (i in 1:n) {  
    Z[i] ~ dnorm(psi[i], tau)I(t.cen[i],)  
  }  
  psi[1:n] <- jump.lin.pred(X[1:n, 1:Q], k, beta.prec)  
  id <- jump.model.id(psi[1:n])  
  beta.prec <- tau / lambda  
  tau ~ dgamma(a, b)  
  k ~ dbin(0.5, Q)  
}  
  
list(  
  n = 3794, Q = 11,  
  lambda = 10000,  
  a = 0.001, b = 0.001,  
  Z = c(0,0,...,12,12), t.cen=c(0,0,...,1,1),  
  X = structure(  
    .Data = c(  
      2,1,1,1,0,0,0,1,1,1.89602,2,  
      1,1,2,1,0,0,0,1,6,2.51399,2,  
      :  
      :  
      1,1,1,0,1,0,0,1,5,4.97648,1,  
      1,1,2,0,1,0,0,1,4,3.19188,1  
    ),  
    .Dim = c(3794, 11))  
)  
  
list(tau = 1)
```


Lampiran 5

Hasil Syntax Reversible Jump MCMC

Hasil

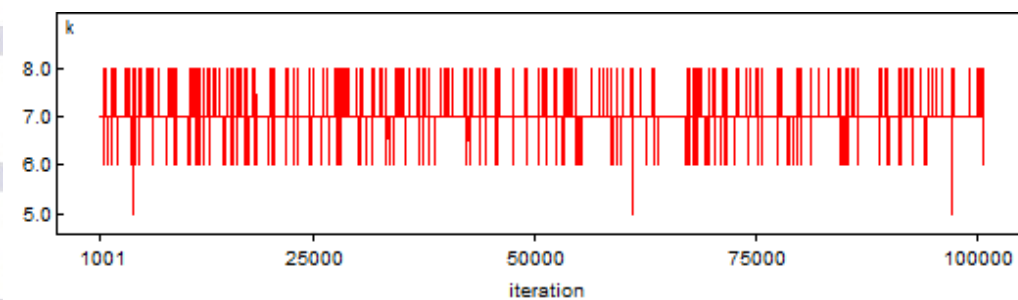
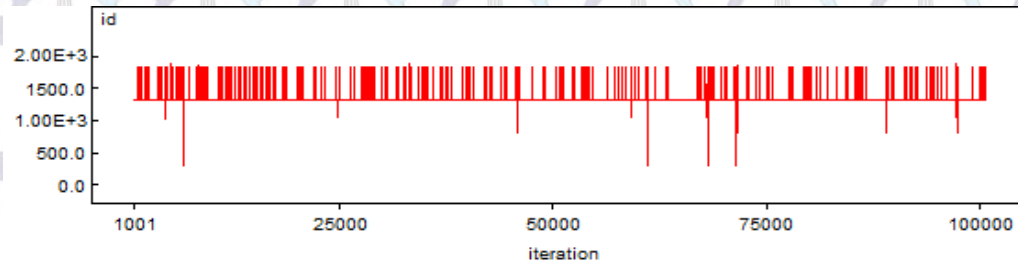
100000 Iterasi

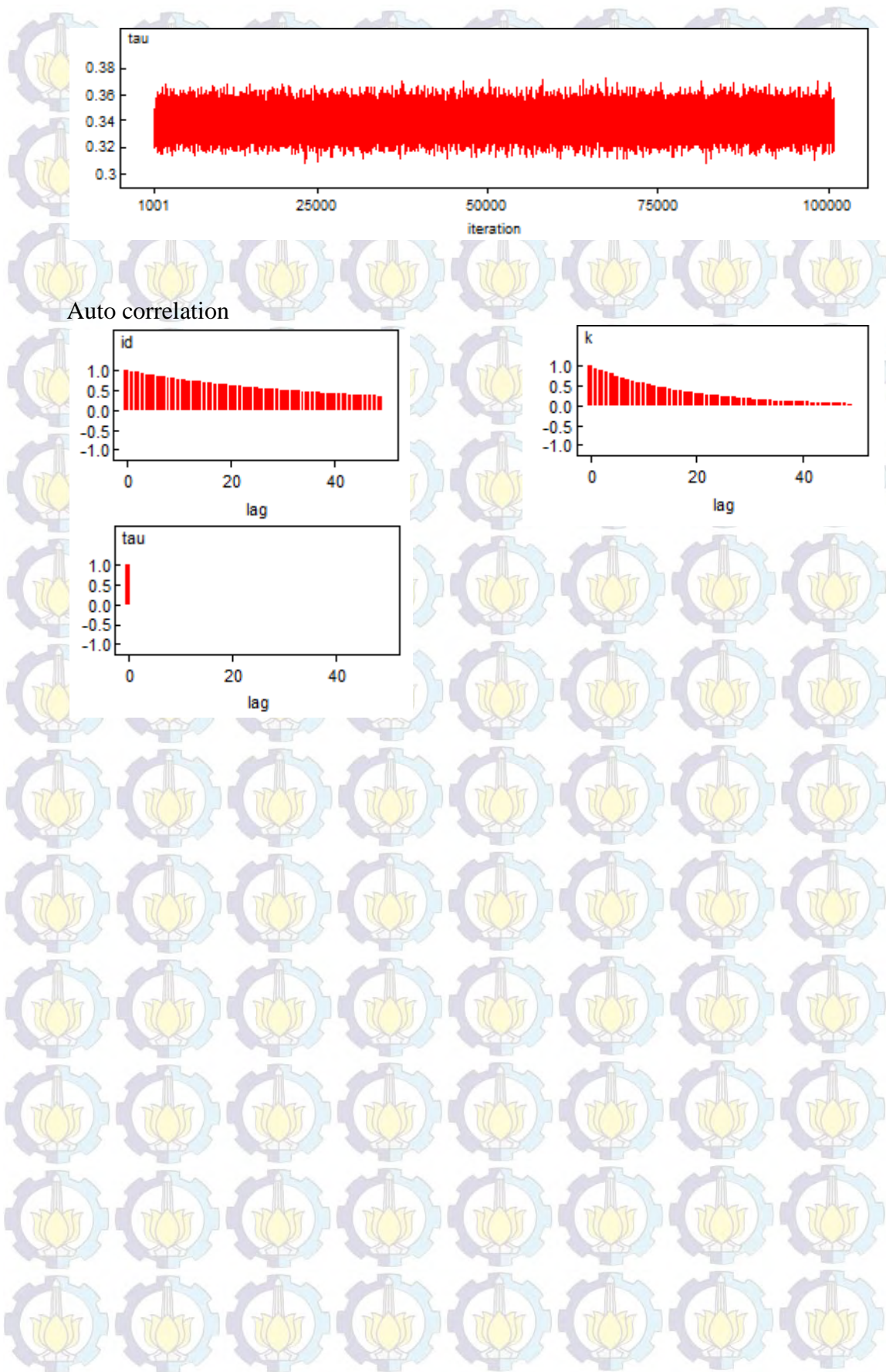
node	mean	sd	MC error 2.5%		median	97.5%	start	sample
id	1335.0	123.2	18.22	1310.0	1311.0	1823.0	1	100000
k	6.012	0.4016	0.06022	6.0	7.0	8.0	1	100000
tau	0.3387	0.01649	7.269E-4	0.3251	0.3393	0.3541	1	100000

model structure	posterior prob.	cumulative prob.
11111000101	0.86021	0.86021
11111000111	0.05878	0.91899
11110000111	0.05303	0.97202
11110000101	0.01396	0.98598
01111000101	0.00557	0.99155

variable no.	marginal prob.
1	0.99296
2	0.99998
3	0.99989
4	1.0
5	0.92995
6	6.0E-4
7	9.1E-4
8	1.7E-4
9	0.99616
10	0.11668
11	0.99826

Time Series





Lampiran 6

Syntax Gibbs Sampling Mixture Normal 6 Komponen

```
model;
{
    for( i in 1 : N ) {
        T[i] ~ dnorm(psi[i],tau[i])|(t.cen[i],)
    }
    for( i in 1 : N ) {
        P[i] ~ dcat(Phi[1:6])
    }

    for( i in 1 : N ) {
        tau[i] <- ptau[P[i]]
    }
    for( i in 1 : N ) {
        psi[i] <- exp(b1[P[i]]*X1[i] + b2[P[i]]*X2[i] + b3[P[i]]*X3[i] + b4_1[P[i]]*X4_1[i] +
b4_2[P[i]]*X4_2[i] + b4_3[P[i]]*X4_3[i] + b4_4[P[i]]*X4_4[i] + b5[P[i]]*X5[i] + b6[P[i]]*X6[i] + b7[P[i]]*X7[i]
+ b8[P[i]]*X8[i] + W[kab[i]])
    }

    for( i in 1:sumNumNeigh)
    {weights[i] <- 1}
    W[1:s] ~ car.normal(adj[], weights[], num[], dtau)

    Phi[1:6] ~ ddirch(Alpha[])
    ptau[1] ~ dgamma(0.1,0.1)
    ptau[2] ~ dgamma(0.1,0.1)
    ptau[3] ~ dgamma(0.1,0.1)
    ptau[4] ~ dgamma(0.1,0.1)
    ptau[5] ~ dgamma(0.1,0.1)
    ptau[6] ~ dgamma(0.1,0.1)
    psigma[1] <- 1 / sqrt(ptau[1])
    psigma[2] <- 1 / sqrt(ptau[2])
    psigma[3] <- 1 / sqrt(ptau[3])
    psigma[4] <- 1 / sqrt(ptau[4])
    psigma[5] <- 1 / sqrt(ptau[5])
    psigma[6] <- 1 / sqrt(ptau[6])

    b1[1] ~ dnorm(-0.167,0.4649)
    b2[1] ~ dnorm(0.548,0.8928)
    b3[1] ~ dnorm(-0.625,0.4604)
    b4_1[1] ~ dnorm(-0.284,10.205)
    b4_2[1] ~ dnorm(0.098,10.407)
    b4_3[1] ~ dnorm(-0.188,10.250)
    b4_4[1] ~ dnorm(0,1)
    b5[1] ~ dnorm(-0.111,0.8911)
    b6[1] ~ dnorm(-0.065,0.1502)
    b7[1] ~ dnorm(-0.087,0.1165)
    b8[1] ~ dnorm(-0.296,0.4421)

    b1[2] ~ dnorm(0.431,0.5605)
    b2[2] ~ dnorm(-0.474,0.7803)
    b3[2] ~ dnorm(-0.182,0.5459)
    b4_1[2] ~ dnorm(-0.195,14.142)
    b4_2[2] ~ dnorm(0.789,13.027)
    b4_3[2] ~ dnorm(-1.363,20.335)
    b4_4[2] ~ dnorm(-1.755,17.678)
    b5[2] ~ dnorm(-0.551,0.6868)
    b6[2] ~ dnorm(0.120,0.1431)
    b7[2] ~ dnorm(-0.056,0.1465)
    b8[2] ~ dnorm(0.182,0.5016)
```

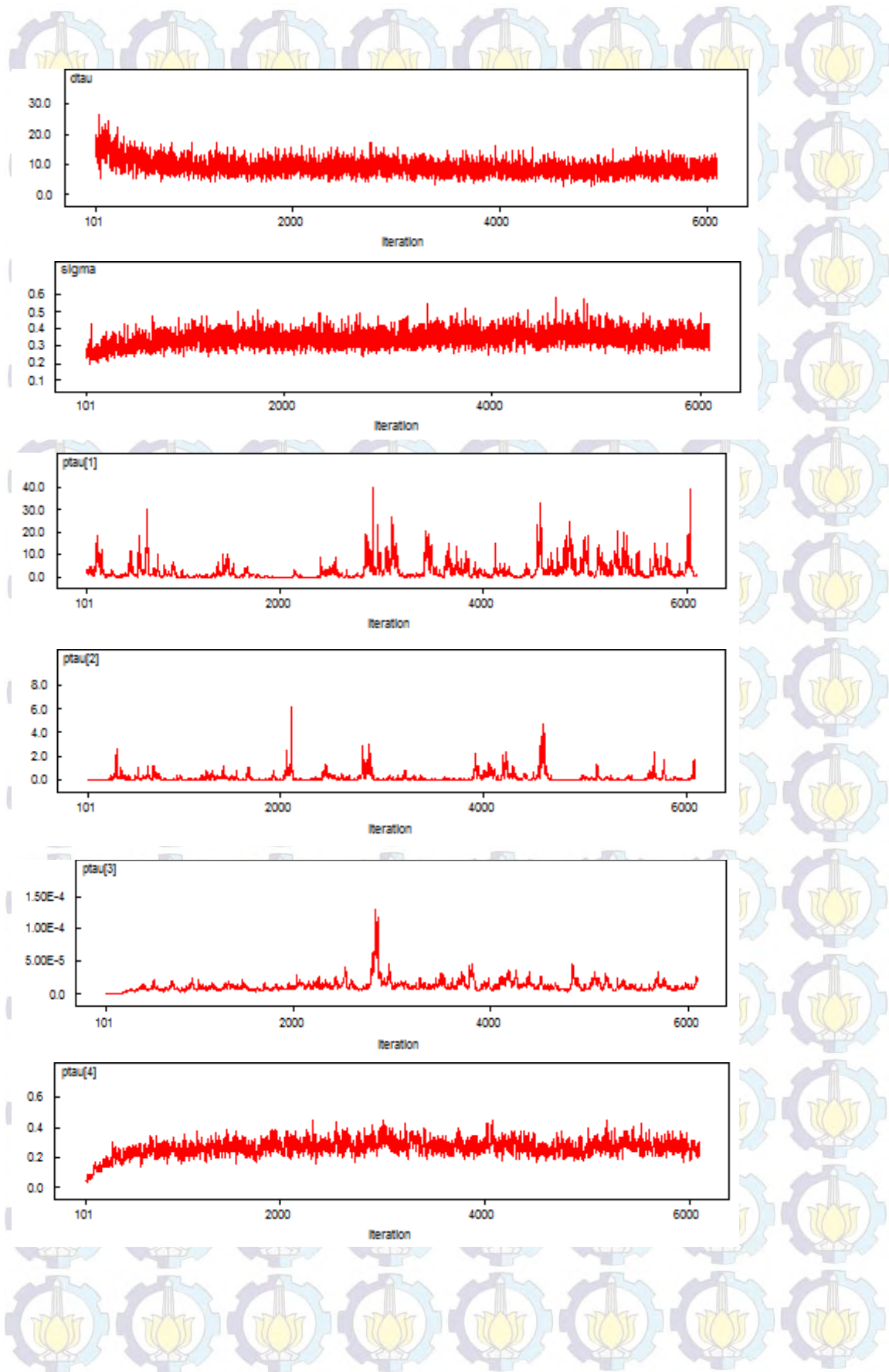

[illegible]

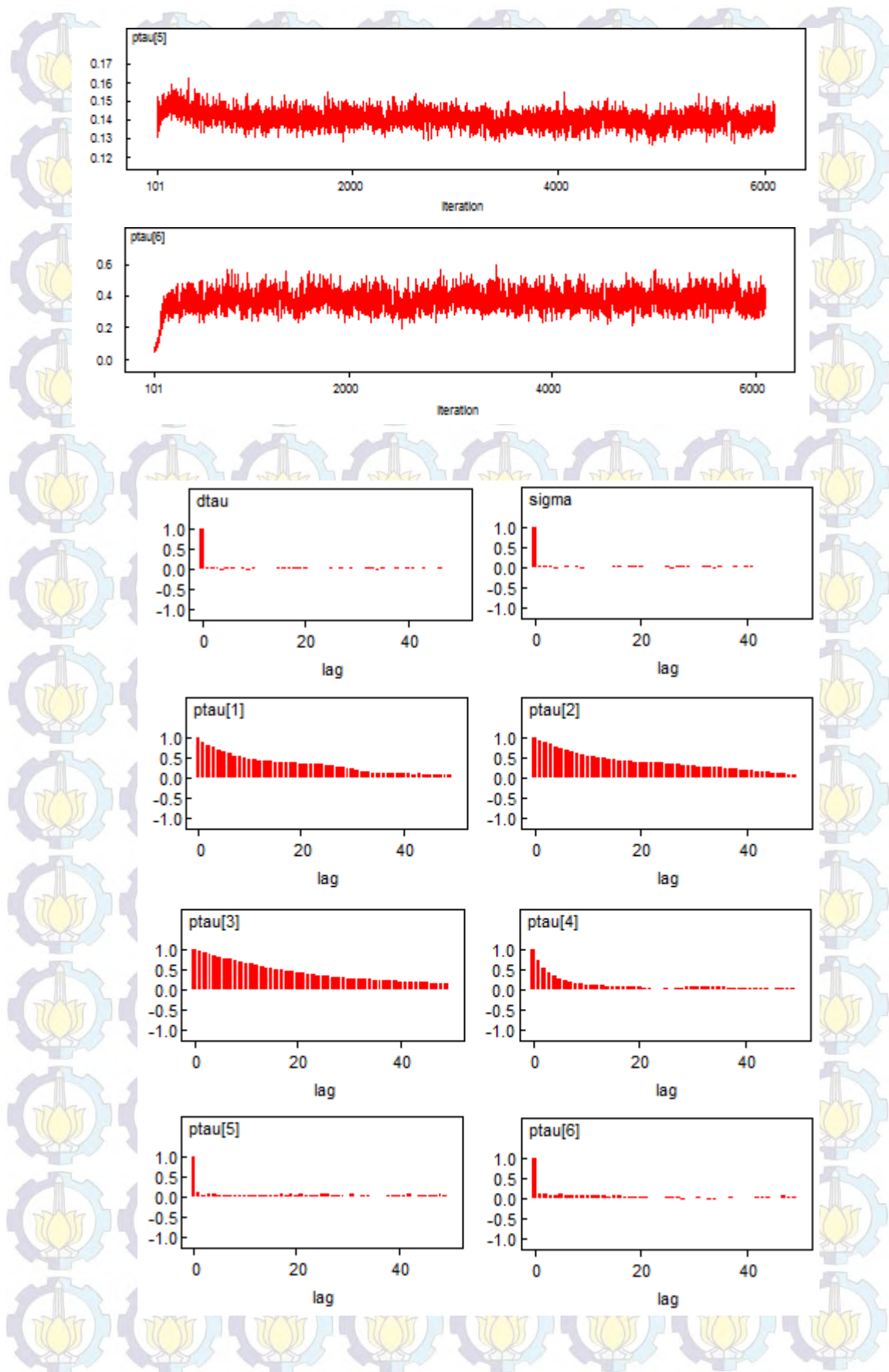
Lampiran 7

Hasil Syntax Gibbs Sampling Mixture Normal 6 Komponen

node	mean sample	sd	MC error	2.5%	median	97.5%	start	
Phi[1]	0.00869	0.001481	3.322E-5	0.005987	0.0086	0.0118	4001	2000
Phi[2]	0.00767	0.001393	2.791E-5	0.005288	0.007587	0.01066	4001	2000
Phi[3]	0.0642	0.004079	9.133E-5	0.05663	0.06416	0.07238	4001	2000
Phi[4]	0.09174	0.004648	8.33E-5	0.08291	0.0916	0.101	4001	2000
Phi[5]	0.7984	0.006549	1.332E-4	0.7856	0.7986	0.8113	4001	2000
Phi[6]	0.02929	0.002773	5.609E-5	0.02411	0.02919	0.03501	4001	2000
W[1]	-0.2717	0.04546	0.006049	-0.3782	-0.2672	-0.1914	4001	2000
W[2]	-0.3483	0.03787	0.005593	-0.4128	-0.3565	-0.285	4001	2000
W[3]	0.00197	0.03312	0.002512	-0.05956	0.001607	0.06699	4001	2000
W[4]	-0.07598	0.02272	0.002929	-0.1185	-0.07568	-0.02693	4001	2000
W[5]	-0.02175	0.03081	0.004051	-0.07908	-0.0228	0.04001	4001	2000
W[6]	-0.009028	0.02444	0.002798	-0.05653	-0.008721	0.0343	4001	2000
W[7]	-0.02994	0.02822	0.003328	-0.07783	-0.03216	0.02657	4001	2000
W[8]	0.05817	0.03094	0.003085	9.871E-4	0.05759	0.1213	4001	2000
W[9]	0.0655	0.02438	0.003243	0.02274	0.065	0.1122	4001	2000
W[10]	-0.01428	0.03373	0.004913	-0.06975	-0.02132	0.06326	4001	2000
W[11]	0.04552	0.03446	0.003978	-0.0161	0.04428	0.1211	4001	2000
W[12]	0.0855	0.03918	0.004739	0.01705	0.08556	0.1652	4001	2000
W[13]	0.00452	0.02288	0.002952	-0.04237	0.004623	0.05193	4001	2000
W[14]	-0.05524	0.03654	0.005401	-0.1101	-0.06866	0.009866	4001	2000
W[15]	0.08199	0.0243	0.001635	0.0331	0.08204	0.1291	4001	2000
W[16]	-0.009999	0.02181	0.002487	-0.05154	-0.01001	0.03211	4001	2000
W[17]	0.0905	0.02432	0.0021	0.04211	0.08982	0.1366	4001	2000
W[18]	-0.0434	0.02463	0.003543	-0.08477	-0.04538	-0.003292	4001	2000
W[19]	0.01097	0.02669	0.001846	-0.04092	0.007773	0.07024	4001	2000
W[20]	0.04101	0.0381	0.002447	-0.03863	0.04283	0.1117	4001	2000
W[21]	0.03902	0.02956	0.00267	-0.01865	0.03782	0.096	4001	2000
W[22]	-0.2173	0.0168	0.002384	-0.249	-0.2142	-0.1918	4001	2000
W[23]	-0.01373	0.0255	0.002468	-0.06406	-0.01291	0.03949	4001	2000
W[24]	-0.1714	0.02835	0.004137	-0.2106	-0.1778	-0.1147	4001	2000
W[25]	0.04598	0.02255	0.001923	0.001085	0.04582	0.09086	4001	2000
W[26]	-0.08357	0.03369	0.00477	-0.1602	-0.08514	-0.023	4001	2000
W[27]	-0.1708	0.02792	0.003044	-0.2244	-0.1722	-0.1075	4001	2000
W[28]	-0.06221	0.02883	0.002299	-0.1202	-0.06231	-0.006239	4001	2000
W[29]	0.03104	0.03067	0.001958	-0.02994	0.03154	0.09379	4001	2000
W[30]	0.137	0.03376	0.00326	0.07085	0.1384	0.2001	4001	2000
W[31]	0.1149	0.03634	0.001972	0.04691	0.1154	0.1807	4001	2000
W[32]	0.1578	0.02699	0.003622	0.1054	0.1563	0.2083	4001	2000
W[33]	0.1685	0.0288	0.001633	0.1138	0.1675	0.2249	4001	2000
W[34]	0.1268	0.02978	0.00356	0.07493	0.1266	0.1816	4001	2000
W[35]	0.1157	0.03315	0.003548	0.05803	0.1146	0.1899	4001	2000
W[36]	0.12	0.0342	0.002553	0.04685	0.1217	0.1847	4001	2000
W[37]	0.07523	0.02374	0.00184	0.03056	0.07608	0.1176	4001	2000
W[38]	-0.01891	0.03777	0.003184	-0.09832	-0.02053	0.05435	4001	2000
b1[1]	0.02705	0.08932	0.01272	-0.1669	0.02487	0.1964	4001	2000
b1[2]	0.5267	0.02618	0.003912	0.479	0.5311	0.5612	4001	2000
b1[3]	0.5389	0.01158	0.001692	0.5177	0.538	0.5625	4001	2000
b1[4]	0.09722	0.03892	0.005682	0.02653	0.09561	0.1833	4001	2000
b1[5]	0.1418	0.01107	0.001421	0.1194	0.1418	0.1627	4001	2000
b1[6]	0.1225	0.02813	0.003531	0.05828	0.124	0.1707	4001	2000
b2[1]	0.9876	0.1508	0.02214	0.7003	0.9863	1.288	4001	2000
b2[2]	0.3795	0.01732	0.002565	0.3456	0.3817	0.4089	4001	2000
b2[3]	0.2462	0.01038	0.001531	0.2274	0.2447	0.2639	4001	2000
b2[4]	0.346	0.05114	0.00753	0.2371	0.3584	0.4167	4001	2000
b2[5]	0.2348	0.01508	0.001908	0.203	0.237	0.2607	4001	2000
b2[6]	0.7286	0.09177	0.01305	0.5384	0.7199	0.912	4001	2000

b3[1]	1.328	0.1621	0.02415	1.091	1.32	1.599	4001	2000
b3[2]	-1.028	0.0147	0.002109	-1.051	-1.032	-0.9999	4001	2000
b3[3]	0.1697	0.01448	0.002136	0.1403	0.1742	0.1933	4001	2000
b3[4]	0.05621	0.03718	0.005375	0.001356	0.04833	0.1394	4001	2000
b3[5]	0.1695	0.009682	0.00108	0.1528	0.1687	0.1882	4001	2000
b3[6]	0.2573	0.02996	0.003523	0.2015	0.2576	0.311	4001	2000
b4_1[1]	-0.08707	0.1801	0.02625	-0.5033	-0.07932	0.2023	4001	2000
b4_1[2]	0.02426	0.07201	0.0107	-0.09264	0.04424	0.1283	4001	2000
b4_1[3]	0.01034	0.01926	0.002817	-0.01658	0.004454	0.04662	4001	2000
b4_1[4]	1.294	0.1004	0.01461	1.136	1.272	1.492	4001	2000
b4_1[5]	0.603	0.0316	0.004293	0.5353	0.602	0.6664	4001	2000
b4_1[6]	-0.04943	0.08181	0.01005	-0.2002	-0.04617	0.1035	4001	2000
b4_2[1]	0.07857	0.3273	0.04377	-0.6468	0.07265	0.7583	4001	2000
b4_2[2]	0.6229	0.0595	0.008888	0.5324	0.613	0.7132	4001	2000
b4_2[3]	0.2694	0.02044	0.003007	0.23	0.2701	0.301	4001	2000
b4_2[4]	1.262	0.1105	0.01595	1.089	1.232	1.515	4001	2000
b4_2[5]	0.6197	0.03245	0.004465	0.5482	0.6193	0.691	4001	2000
b4_2[6]	0.153	0.06349	0.007874	0.03974	0.1538	0.2768	4001	2000
b4_3[1]	-0.1805	0.2108	0.03023	-0.5683	-0.1531	0.1717	4001	2000
b4_3[2]	-1.329	0.2244	0.02558	-1.825	-1.32	-0.8784	4001	2000
b4_3[3]	-1.024	0.1288	0.01812	-1.266	-1.021	-0.7419	4001	2000
b4_3[4]	1.375	0.1336	0.01912	1.149	1.383	1.617	4001	2000
b4_3[5]	0.67	0.0315	0.004173	0.6084	0.6691	0.7387	4001	2000
b4_3[6]	0.09768	0.06401	0.007685	-0.02007	0.09498	0.2263	4001	2000
b4_4[1]	-0.08752	0.3676	0.0499	-0.7417	-0.1116	0.7962	4001	2000
b4_4[2]	-1.766	0.2069	0.02584	-2.241	-1.768	-1.43	4001	2000
b4_4[3]	0.1204	0.3459	0.05139	-0.5799	0.2927	0.5987	4001	2000
b4_4[4]	1.302	0.1066	0.01574	1.103	1.298	1.514	4001	2000
b4_4[5]	0.69	0.02999	0.003995	0.6328	0.69	0.7594	4001	2000
b4_4[6]	0.05246	0.05544	0.006933	-0.04913	0.04924	0.1588	4001	2000
b5[1]	0.4103	0.09288	0.0131	0.2095	0.408	0.5642	4001	2000
b5[2]	-0.8042	0.02204	0.003236	-0.8425	-0.803	-0.769	4001	2000
b5[3]	-0.2245	0.01666	0.00247	-0.2509	-0.2278	-0.1948	4001	2000
b5[4]	0.1496	0.08143	0.01203	0.0139	0.1631	0.2691	4001	2000
b5[5]	0.2619	0.0141	0.001771	0.2347	0.2604	0.2922	4001	2000
b5[6]	0.1959	0.03887	0.004955	0.1236	0.1953	0.2728	4001	2000
b6[1]	0.003723	0.02156	0.002909	-0.03822	8.446E-4	0.05784	4001	2000
b6[2]	0.1589	0.005748	8.439E-4	0.1485	0.16	0.1685	4001	2000
b6[3]	0.5691	0.002693	4.007E-4	0.5652	0.5688	0.5729	4001	2000
b6[4]	0.07035	0.002363	3.496E-4	0.06517	0.07105	0.07355	4001	2000
b6[5]	0.01332	5.71E-4	3.927E-5	0.01206	0.01334	0.01443	4001	2000
b6[6]	0.004926	0.002667	2.709E-4	-8.163E-4	0.004963	0.009834	4001	2000
b7[1]	-0.02557	0.02023	0.002779	-0.07475	-0.02323	0.003989	4001	2000
b7[2]	0.7912	0.002474	3.614E-4	0.7861	0.7922	0.7945	4001	2000
b7[3]	0.3746	0.003618	5.386E-4	0.3694	0.3743	0.3811	4001	2000
b7[4]	0.02315	0.01196	0.001731	0.001777	0.02325	0.04526	4001	2000
b7[5]	0.07189	0.003826	5.083E-4	0.06466	0.07165	0.07988	4001	2000
b7[6]	0.05403	0.0124	0.001632	0.02615	0.05466	0.0779	4001	2000
b8[1]	-0.8775	0.08503	0.01228	-1.015	-0.8855	-0.6988	4001	2000
b8[2]	0.5496	0.00953	0.001406	0.5362	0.5468	0.5681	4001	2000
b8[3]	0.01988	0.004412	6.321E-4	0.01068	0.02018	0.02699	4001	2000
b8[4]	0.0706	0.03307	0.004793	0.01068	0.06714	0.1427	4001	2000
b8[5]	0.2039	0.0107	0.001354	0.1821	0.2045	0.2215	4001	2000
b8[6]	0.5175	0.062	0.008505	0.3849	0.5171	0.6355	4001	2000
ptau[1]	5.404	5.85	0.6096	0.2727	3.287	21.5	4001	2000
ptau[2]	0.3685	0.5288	0.0524	0.009969	0.181	1.925	4001	2000
ptau[3]	0.04391	0.05024	0.006787	0.004696	0.0251	0.1833	4001	2000
ptau[4]	0.2754	0.04109	0.003276	0.2057	0.2728	0.3656	4001	2000
ptau[5]	0.1444	0.003917	1.259E-4	0.1365	0.1445	0.152	4001	2000
ptau[6]	0.3812	0.05655	0.002819	0.2763	0.38	0.4968	4001	2000





MODEL *MIXTURE SURVIVAL* SPASIAL PADA ANGKA LAMA SEKOLAH ANAK UMUR 16-18 TAHUN DI JAWA TIMUR TAHUN 2012

¹Jihadil Qudsi S. dan ²Nur Iriawan

^{1,2}Statistika, FMIPA, Institut Teknologi Sepuluh Nopember (ITS)

Jl. Arief Rahman Hakim, Surabaya 60111 Indonesia

e-mail: ¹jihadil.qudsi@gmail.com dan ²nuririawan@gmail.com

ABSTRAK

Analisis *survival* merupakan metode data lama waktu suatu obyek sampai terjadinya peristiwa tertentu yang disebut sebagai *failure event*. Perkembangan metode ini cukup pesat, bahkan telah sukses dipadukan dengan efek spasial serta pemberian konsep *mixture* untuk diterapkan pada kasus spasial. Penelitian ini bertujuan untuk memperoleh model *survival* yang mengakomodasi konsep *mixture* dengan efek spasial untuk menentukan variabel-variabel yang berpengaruh dalam model.

Model ini diterapkan pada data angka lama sekolah anak umur 16-18 tahun se Jawa Timur. Data tersebut merupakan data sekunder dari hasil survei sosial ekonomi nasional (susenas) tahun 2012 oleh Badan Pusat Statistik Provinsi Jawa Timur. Variabel-variabel yang diperkirakan mempengaruhi angka lama sekolah anak umur 16-18 tahun yaitu jenis kelamin, status perkawinan, status bekerja, tingkat pendidikan kepala rumah tangga, jenis kelamin kepala rumah tangga, rata-rata pengeluaran rumah tangga per kapita, jumlah anggota rumah tangga, dan lokasi tempat tinggal (kota atau desa). Penelitian ini menentukan banyak komponen *mixture* dengan algoritma RJMCMC, lalu dilakukan estimasi parameter untuk mencari variabel-variabel yang secara signifikan mempengaruhi angka lama sekolah anak umur 16-18 tahun dengan Gibbs Sampling menggunakan software WinBUGS.

Berdasarkan hasil RJMCMC yang paling optimum, didapat 6 komponen *mixture* pada data angka lama sekolah anak umur 16-18 tahun dengan variabel-variabel prediktornya. Sedangkan hasil dari *Gibbs Sampling* didapat bahwa pada setiap komponen *mixture*, variabel prediktor yang secara signifikan mempengaruhi variabel respon berbeda-beda dan efek spasial dari daerah kabupaten/kota tidak semua secara signifikan mempengaruhi angka lama sekolah anak umur 16-18 tahun.

Kata Kunci: Analisis *Survival*, Angka Lama Sekolah, *Mixture Survival* Spasial RJMCMC.

PENDAHULUAN

Latar Belakang

Analisis *survival* adalah metode statistika yang mempelajari lamanya suatu kejadian atau peristiwa yang terjadi dan biasa dikenal dengan nama *failure event*. Waktu ketahanan hidup (*survival time*) atau *T* merupakan waktu dari awal perlakuan sampai terjadinya respon pertama kali yang ingin diamati. Respon yang dimaksud berupa waktu yang diperlukan hingga suatu peristiwa yang diharapkan terjadi dan mungkin saja belum ditemukan pada saat pengumpulan data berakhir sehingga waktu *survival*-nya tidak dapat diamati. Pada kondisi demikian, pengamatan disebut sebagai pengamatan tersensor [1]. Metode regresi *survival* adalah metode regresi yang digunakan untuk melihat variabel-variabel yang menyebabkan terjadinya suatu peristiwa (biasa dikenal dengan nama *time dependent covariate*) dengan variabel responnya adalah waktu ketahanan hidup. Salah satu metode regresi *survival* yang sering digunakan adalah regresi *cox proportional hazard*.

Perkembangan dari analisis *survival* dengan menambahkan konsep distribusi *mixture*. Distribusi *mixture* menggabungkan sejumlah komponen yang kemungkinan berasal dari distribusi yang sama atau bahkan berbeda sehingga dapat memberikan gambaran yang lebih spesifik mengenai sifat-sifat data. Model yang menggunakan konsep distribusi *mixture* akan memberikan hasil yang lebih akurat karena lebih mendekati distribusi data yang sebenarnya [2].

Pengembangan lain dari analisis *survival* dengan penyertaan efek spasial pada analisis *survival*, dimana efek

spasial seringkali mempengaruhi waktu terjadinya suatu kejadian. Penyertaan efek spasial pada analisis *survival* oleh beberapa peneliti menambahkan model efek random untuk mengatasi heterogenitas/sumber-sumber variansi yang tidak ter jelaskan dalam model [3].

Beberapa penelitian diantaranya oleh Banerjee dan Carlin [4] yang menggunakan metode *survival* dengan efek spasial frailty pada data kematian di Minnesota, penelitian tersebut membandingkan 4 model yaitu model *survival* non frailty, model *survival* non-spasial frailty, model *survival* frailty CAR, dan model *survival* frailty geostat, hasilnya model *survival* frailty CAR yang terbaik. Lain halnya dengan penelitian oleh Aksioma [5] menggunakan data kejadian HIV/AIDS dengan metode *survival* spasial, dan penelitian oleh Hasyim [6] yang menambahkan efek spasial dan konsep *mixture* pada metode *survival* dan diterapkan pada data penderita demam berdarah.

Analisis *survival* sering digunakan pada bidang kesehatan, teknik, industri, sosiologi dan ekonomi. Beberapa penelitian analisis *survival* di atas juga lebih banyak meneliti kasus pada bidang kesehatan. Dalam penelitian ini peneliti mencoba menggunakan analisis *survival* pada bidang sosiologi yang mencakup bidang pendidikan.

Indikator pendidikan merupakan salah satu indikator dalam membangun Indeks Pembangunan Manusia (IPM), dimana ada dua angka yang digunakan dalam membangun indikator pendidikan yaitu angka persentase warga melek huruf dan angka rata-rata lama sekolah. Guna meningkatkan rata-rata lama sekolah, pemerintah memberlakukan wajib belajar 9 tahun bagi warga setiap warga yang mengenyam

pendidikan. Program ini mewajibkan setiap warga negara Indonesia untuk bersekolah selama 9 (sembilan) tahun pada jenjang pendidikan dasar, yaitu dari Sekolah Dasar (SD) hingga Sekolah Menengah Pertama (SMP).

IPM provinsi Jawa Timur tahun 2012 oleh Badan Pusat Statistik Jawa Timur, dari 38 kabupaten/kota di provinsi Jawa Timur hanya 9 kabupaten/kota yang memiliki rata-rata lama sekolah di atas 9 tahun, artinya sebagian besar warga di 29 kabupaten/kota lainnya di provinsi Jawa Timur tidak memiliki ijazah SMP/ sederajat yang formal. Pada tahun yang sama, tingkat partisipasi sekolah anak umur 16-18 tahun (dalam usia sekolah menengah atas) di Jawa Timur sebesar 61,87% dari total anak yang berumur 16-18 tahun yang dapat bersekolah [7]. Jika kita melihat pada 3 tahun sebelumnya, tahun 2009, dimana anak yang berumur 16-18 tahun pada tahun 2012, berumur 13-15 tahun (dalam usia sekolah menengah pertama), tingkat partisipasi sekolahnya sebesar 88%. Berarti dalam waktu 3 tahun terjadi penurunan tingkat partisipasi sekolah sebesar 26,13% dari jenjang umur 13-15 ke jenjang umur 16-18 tahun.

Fenomena banyaknya anak yang putus sekolah karena himpitan ekonomi sehingga harus berhenti sekolah karena bekerja atau bahkan *drop out* (dikeluarkan dari sekolah). Selain itu bisa terjadi karena kapasitas daya tampung murid terbatas yang disediakan oleh pemerintah disetiap jenjang pendidikan. Dilihat pada rasio murid terhadap banyak sekolah menengah pertama dan sederajat yang berstatus negeri dan swasta di Jawa Timur pada tahun ajaran 2012/2013 yaitu 270/1, artinya secara umum 1 sekolah hanya bisa menampung maksimal 270 anak setiap tahun ajaran baru. Banyak sekolah negeri dan swasta yang ada di Jawa Timur sebanyak 6.824 sekolah, hal tersebut berarti hanya 1.839.607 anak yang dapat ditampung untuk bersekolah pada tingkat menengah pertama atau yang sederajat setiap ajaran baru, sedangkan banyak anak yang berumur 13-15 tahun pada tahun 2012 sekitar 2.007.866 anak [7].

Penelitian memodelkan angka lama sekolah sendiri sudah pernah diteliti oleh Santoso [8] dengan menggunakan metode pendekatan *spline multivariable* dan MARS, dan juga oleh Sulistyawati [9] yang menerapkan metode analisis *survival* dengan pemberian konsep *mixture*.

Efek spasial dalam penelitian ini dikarenakan daerah yang memiliki jumlah sekolah dan fasilitas yang memadai akan cenderung lebih diminati oleh penduduk daerah lain untuk menyekolahkan anak mereka ke daerah tersebut. Kedekatan antar daerah dengan angka rata-rata lama sekolah tinggi atau rendah juga diperkirakan akan memperkuat dugaan adanya efek spasial. Maka berdasarkan uraian di atas judul penelitian yang diajukan peneliti yaitu model *mixture survival* spasial pada angka lama sekolah anak umur 16-18 tahun di Jawa Timur tahun 2012.

Rumusan Masalah

Adapun permasalahan yang diangkat dalam penelitian ini adalah

1. Bagaimana model *survival* dengan penambahan konsep *mixture* dan pertimbangan adanya efek spasial menggunakan pendekatan bayesian?
2. Bagaimana model *mixture survival* spasial pada angka lama sekolah anak umur 16-18 tahun di Jawa Timur berdasarkan variabel-variabel yang mempengaruhi tinggi atau rendahnya angka lama sekolah?

Tujuan Penelitian

Berdasarkan rumusan masalah diatas, maka tujuan yang ingin dicapai dalam penelitian ini adalah :

1. Mendapatkan model *survival* dengan penambahan konsep *mixture* dan pertimbangan adanya efek spasial menggunakan pendekatan bayesian
2. Mendapatkan model *mixture survival* spasial pada angka lama sekolah anak umur 16-18 tahun di Jawa Timur berdasarkan variabel-variabel yang mempengaruhi tinggi atau rendahnya angka lama sekolah.

TINJAUAN PUSTAKA

Definisi Analisis Survival

Analisis *survival* adalah salah satu metode statistika untuk menganalisis data dimana variabel responnya berupa waktu sampai suatu *event* terjadi. *Event* didefinisikan sebagai peristiwa ekstrim yang mungkin terjadi pada individu, misalnya kesembuhan maupun kematian setelah dilakukan proses pengobatan. Sedangkan respon *survival* didefinisikan sebagai *range* waktu dari awal penelitian sampai suatu *event* terjadi atau sampai penelitian berakhir, misalnya *range* waktu awal penelitian sampai individu mati atau sampai penelitian berakhir [10].

Ada tiga elemen yang harus diperhatikan dalam menentukan waktu *survival* t [11], *time origin/starting point*, *failure time*, dan *measurement scale of time*. Dalam menganalisis dengan tipe data lama waktu tidak dapat menggunakan analisis regresi linier biasa karena terdapat data tersensor dan tidak pula dapat menggunakan analisis regresi logistik karena lamanya waktu data yang diikuti tidak sama, sehingga yang paling sesuai adalah menggunakan analisis *survival* [12]. Tiga alasan terjadinya penyensoran [13], yaitu *lost of follow up*, *drop out*, dan *termination of study*.

Model *survival* digunakan untuk menjelaskan bagaimana resiko (*hazard*) terjadinya suatu *event* tertentu pada suatu waktu dipengaruhi oleh beberapa *covariate* berdasarkan teori yang menunjang peristiwa tersebut.

Fungsi Survival dan Fungsi Hazard

Misalkan T adalah variabel acak non negatif yang menggambarkan waktu *survival* individu dari suatu populasi. Peluang T pada analisis *survival* secara umum digambarkan ke dalam tiga fungsi yaitu fungsi kepadatan peluang, fungsi distribusi kumulatif, fungsi *survival*, dan fungsi *hazard*. Fungsi *survival* $S(t)$ dapat dinyatakan sebagai peluang seorang individu untuk bertahan lebih lama dari suatu waktu t dan dinyatakan melalui persamaan berikut.

$$S(t) = P(T > t) = 1 - P(T \leq t) = 1 - F(t) \quad (1)$$

Sedangkan fungsi *hazard* dijelaskan sebagai reaksi sesaat atau laju kegagalan (*failure*) sesaat ketika mengalami suatu *event* pada waktu ke- t dan dinyatakan sebagai berikut.

$$h(t) = \lim_{\Delta t \rightarrow 0} \left\{ \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t} \right\} \quad (2)$$

Kleinbaum & Klein (2005) menyatakan hubungan antara fungsi *survival* dan fungsi *hazard* sebagai berikut:

$$H(t) = -\ln S(t) \quad (3)$$

dengan : t : Waktu yang diamati

T : Waktu *survival* seorang individu

$F(t)$: Fungsi distribusi kumulatif dari distribusi data

$S(t)$: Fungsi *survival*

$h(t)$: Fungsi *hazard* proporsional

Model dan Asumsi Hazard Proporsional

Nilai variabel prediktor pada model *hazard* proporsional dinyatakan oleh vektor \mathbf{x} , dimana $\mathbf{x} = (x_1, x_2, \dots, x_p)$. Fungsi *baseline hazard* dinyatakan sebagai $h_0(t)$ merupakan fungsi *hazard* untuk tiap-tiap individu dimana semua variabel prediktor dalam vektor \mathbf{x} bernilai 0 [14]. Secara umum, model *hazard* proporsional diberikan dalam persamaan (4) berikut,

$$h(t, \mathbf{x}) = h_0(t) \exp(\beta^T \mathbf{x}) \quad (4)$$

atau dapat dituliskan

$$h(t, \mathbf{x}) = h_0(t) \exp(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p) \quad (5)$$

Persamaan (4) disebut juga sebagai regresi *cox*.

Hal yang menarik pada regresi *cox* adalah data tidak harus memenuhi distribusi apapun [15]. Sedangkan asumsi pemodelan yang harus dipenuhi dalam regresi *cox* yaitu asumsi *hazard* proporsional yang berarti fungsi *hazard* harus proporsional setiap waktu karena regresi *cox* tidak mengakomodasi variabel yang berubah-ubah sepanjang waktu [14]. Proporsional berarti bahwa variabel prediktor independen terhadap waktu dan hubungan antara *hazard* kumulatif sudah proporsional setiap waktu. Asumsi proporsional tersebut dapat diketahui dengan melihat pola plot $-\ln[-\ln S(t)]$ atau $\ln[-\ln S(t)]$ terhadap waktu *survival* untuk setiap variabel prediktor dengan skala kategorik. Asumsi *hazard* proporsional terpenuhi jika pola plot antar kategori dalam variabel prediktor membentuk pola yang sejajar [10]. Sedangkan untuk variabel prediktor dengan skala ratio tidak memiliki asumsi apapun.

Model Survival Spasial

Banerjee *et al.* [4] mengelompokkan data waktu hingga terjadinya suatu peristiwa (*time-to-event data*) ke dalam strata-strata/kelompok-kelompok, seperti wilayah geografis atau daerah bencana. Hal diperkenalkan oleh Banerjee *et al.* [4] dimana terdapat *mixed model* dengan efek random yang mewakili status kesehatan tiap kelompok.

Misalkan t_{ij} menyatakan waktu hingga terjadinya suatu *event* atau waktu hingga suatu individu i (dimana $i = 1, 2, \dots, n$) dalam strata/kelompok j (dimana $j = 1, 2, \dots, m$) mengalami kejadian lulus, sedangkan \mathbf{x}_{ij} menyatakan vektor dari *covariate* yang mempengaruhi kejadian tersebut, maka asumsi *hazard* proporsional $h(t_{ij}; \mathbf{x}_{ij})$ memungkinkan untuk membentuk model semiparametrik *cox* dengan *hazard rate*,

$$h(t_{ij}; \mathbf{x}_{ij}) = h_0(t_{ij}) \exp(\beta^T \mathbf{x}_{ij}) \quad (6)$$

dimana *baseline hazards* (h_0) hanya dipengaruhi oleh hasil perkalian dari bentuk eksponensial yang menyertakan *covariate*.

Model *Cox* tersebut tidak memuat *intercept* karena tidak dilakukan pengukuran parameter *baseline hazard* [3]. Dalam model yang menyertakan *frailty*, maka persamaan (6) kemudian dapat diperluas menjadi,

$$h(t_{ij}; \mathbf{x}_{ij}) = h_0(t_{ij}) \exp(\beta^T \mathbf{x}_{ij} + W_j) \quad (7)$$

dimana W_j merupakan bentuk *stratum-specific frailty* yang dibentuk untuk menyatakan perbedaan antar strata. Secara umum W_j diasumsikan sebagai berikut,

$$W_j \stackrel{iid}{\sim} N(0, \sigma^2)$$

Menurut [4], distribusi non-normal juga sering digunakan untuk memodelkan *frailties*. Artinya tidak harus distribusi normal untuk asumsi W_j .

Model pada pendekatan spasial *survival* dibentuk dari data *survival* yang tersusun secara spasial, artinya *frailties* W_j dari kelompok/strata yang saling berdekatan menggambarkan kemungkinan bahwa kelompok/strata tersebut memiliki nilai/karakteristik yang serupa, misalnya kelompok kabupaten

di sebuah provinsi [3][4]. Susunan spasial dari kelompok-kelompok tersebut dapat dimodelkan dalam dua cara yaitu pendekatan geostatistik dan pendekatan *lattice*.

Model Lattice

Model ini menggunakan informasi jarak dari ketetanggaan (*neighborhood*) atau kedekatan antara satu *region* dengan *region* yang lain. Model ini menggunakan metode penggabungan informasi tentang wilayah-wilayah yang saling berdekatan/ber-tetangga dibandingkan dengan informasi jarak metrik [4]. Distribusi dari efek random \mathbf{W}^* didefinisikan sebagai $\mathbf{W}^* | \lambda \sim \text{CAR}(\lambda)$, yaitu model *conditionally autoregressive* yang menyatakan adanya ketergantungan spasial, dimana $\lambda = 1/v$ yang merupakan parameter distribusi *CAR* yang menyatakan presisi atau *inverse* dari varians distribusi efek randomnya [16].

Matriks Penimbang Spasial

Dalam konteks ketetanggaan (*lattice*) ini, matriks penimbang spasial dari daerah-daerah yang saling bersinggungan dinyatakan melalui indeks diskret [4]. Beberapa metode yang mendefinisikan hubungan kebersinggungan (*contiguity*) antar daerah menurut LeSage [17] antara lain: *Linear Contiguity* (persinggungan tepi), *Rook Contiguity* (persinggungan sisi), *Bishop Contiguity* (persinggungan sudut), *Double Linear Contiguity* (persinggungan dua tepi), *Double Rook Contiguity* (persinggungan dua sisi), *Queen Contiguity* (persinggungan sisi sudut).

Model Mixture

Model *mixture* akan tampak dari data yang diamati dan data yang ada biasanya terdiri dari beberapa subpopulasi atau grup [18]. Setiap subpopulasi merupakan suatu komponen dari *mixture* tertentu dengan proporsi yang bervariasi untuk setiap komponennya. Distribusi *mixture* menggabungkan sejumlah komponen yang kemungkinan berasal dari distribusi yang sama atau bahkan berbeda-beda sehingga dapat memberikan gambaran mengenai sifat-sifat data. Hasil distribusi *mixture* dapat memfasilitasi deskripsi suatu sistem yang kompleks dengan lebih teliti. Untuk mendeteksi kecenderungan bahwa suatu data berdistribusi *mixture* dapat dilihat dari histogram data tersebut [19].

Misalkan terdapat k komponen dalam sebuah *mixture*, diberikan $f_1(x), f_2(x), \dots, f_K(x)$ sebagai komponen densitas *survival* pertama, kedua sampai K komponen, maka model *mixture* yang disusun oleh k komponen tersebut akan dapat dituliskan :

$$f(x | \boldsymbol{\pi}) = \pi_1 f_1(x) + \pi_2 f_2(x) + \dots + \pi_K f_K(x) \quad (8)$$

dengan π_1 : Nilai proporsi dari komponen *mixture* pertama

π_2 : Nilai proporsi dari komponen *mixture* kedua

π_K : Nilai proporsi dari komponen *mixture* ke- K

sehingga $\pi_1 + \pi_2 + \dots + \pi_K = 1$

Model Frailty

Model semiparametrik *Cox* yang telah dijelaskan sebelumnya mengasumsikan bahwa faktor-faktor yang mempengaruhi *hazard* suatu kejadian telah dijelaskan dalam vektor *covariate* \mathbf{x} . Pada kenyataannya terdapat kasus-kasus tertentu (misalnya kasus dengan data spasial) dimana terdapat keragaman/sumber-sumber variansi yang tidak dapat dijelaskan melalui vektor *covariate* dalam model. Akibatnya, terjadi bias pada pendugaan parameter *survival*. Salah satu cara untuk mengatasi permasalahan tersebut yaitu dengan menyertakan efek random atau *frailty term* dalam model [3]. Dalam konteks spasial *survival* digunakan bentuk *frailty*

hirarki atau *stratum-specific frailties* dimana unit pengamatan dikelompokkan dalam wilayah-wilayah *event* yang diteliti.

Analisis Bayesian

Metode Bayesian ditemukan oleh seorang ilmuwan bernama Thomas Bayes (1702-1761), yang menemukan perlakuan matematika untuk masalah *non trivial* dari inferensi Bayesian. Analisis bayesian memperlakukan semua parameter yang tidak diketahui sebagai variabel random dan memiliki distribusi [20]. Teorema Bayesian didasarkan pada distribusi *posterior* yang merupakan perpaduan antara distribusi *prior* (informasi masa lalu sebelum dilakukan observasi) dan data observasi yang digunakan untuk menyusun fungsi *likelihood* [21]. Hubungan distribusi *posterior* dengan distribusi *prior* dan *likelihood* dapat dituliskan sebagai berikut :

Distribusi *posterior* \propto *likelihood* \times Distribusi *prior*

Pada teorema Bayes, parameter θ yang diberikan oleh data observasi y , maka distribusi probabilitas untuk *posterior* θ pada data y akan proporsional dengan perkalian antara distribusi *prior* θ dan fungsi *likelihood* θ yang diberikan oleh data y . Secara matematis dapat ditulis sebagai berikut :

$$f(\theta | x) = \frac{f(x | \theta)f(\theta)}{f(x)} \quad (9)$$

$$f(\theta | x) \propto f(x | \theta)f(\theta)$$

Dimana $f(\theta|y)$ merupakan distribusi *posterior* yang proporsional dengan perkalian antara fungsi *likelihood* $f(x|\theta)$ dan distribusi *prior* $f(\theta)$.

Likelihood Pada Model Mixture

Fungsi *likelihood* distribusi *mixture* berbeda dengan fungsi *likelihood* distribusi univariat biasa [18]. Jika terdapat data pengamatan sebanyak n yang terdekomposisi ke dalam k kelompok data (subpopulasi) yang masing-masing mempunyai distribusi maka fungsi *likelihood* model *mixture* tersusun dari beberapa *likelihood* data di setiap subpopulasi menurut distribusi masing-masing.

$$l_{mix} = \prod_{i=1}^n f_{mix}(x_i | \theta)$$

$$l_{mix} = \prod_{i_1=1}^{n_1} \pi_1 f(x_{i_1} | \theta_1) + \prod_{i_2=1}^{n_2} \pi_2 f(x_{i_2} | \theta_2) + \dots + \prod_{i_k=1}^{n_k} \pi_k f(x_{i_k} | \theta_k) \quad (10)$$

dengan syarat adalah $n_1 + n_2 + \dots + n_k = n$ dan k adalah banyaknya komponen *mixture*.

Distribusi Prior

Distribusi *prior* merupakan informasi yang terdahulu mengenai parameter [21]. Berikut beberapa macam distribusi *prior* dalam metode Bayesian, yaitu *Proper prior* atau *improper prior*, *Informative prior* atau *non informative prior*, *Conjugate prior* atau *non conjugate prior*, dan *Pseudo prior*.

Markov Chain Monte Carlo (MCMC)

Menurut Carlin [22], pendekatan MCMC sangat efektif untuk mengurangi beban komputasi dalam menyelesaikan persamaan integrasi yang kompleks dan metode ini memungkinkan proses simulasi dengan mengambil sampel acak dari model stokastik yang sangat rumit. Ide dasar dari MCMC adalah membangkitkan data sampel dari distribusi *posterior* sesuai proses markov chain dengan menggunakan simulasi monte carlo secara iteratif sehingga diperoleh kondisi yang konvergen terhadap *posterior* [23].

Gibbs Sampling

Gibbs Sampling didefinisikan sebagai suatu teknik simulasi untuk membangkitkan variabel random dari suatu

fungsi distribusi tertentu tanpa harus menghitung fungsi densitasnya. Proses *Gibbs Sampling* dilakukan dengan mengambil sampel dengan cara membangkitkan rangkain *gibbs* variabel random berdasarkan sifat-sifat dasar proses Markov Chain. Hal ini merupakan kelebihan dari *Gibbs Sampling* karena variabel random tersebut dibangkitkan dengan menggunakan konsep distribusi unidimensional yang terstruktur sebagai distribusi full conditional. *Gibbs Sampling* sangat berguna dalam mengestimasi suatu parameter yang mempunyai tingkat kerumitan dalam proses integritasi yang kompleks dan sulit diselesaikan secara analitis.

Reversible Jump MCMC

Pada umumnya penggunaan algoritma MCMC pada proses inferensi bayesian terbatas hanya kasus dimana dimensi vektor parameter model bersifat tetap. Hal ini akan menjadi tidak berlaku pada kondisi dimana dimensi vektor parameter model tidak tetap. Oleh karena itu Green [24] mengembangkan algoritma Metropolis Hasting untuk mengatasi permasalahan tersebut yang kemudian disebut dengan *Reversible Jump Markov Chain Monte Carlo* (RJMCMC).

Pada pemodelan *mixture* dengan banyak komponen penyusun *mixture* yang tidak diketahui. Richardson dan Green [25] memanfaatkan algoritma ini untuk mendapatkan model *mixture* yang tepat bagi data yang terindikasi bersifat multimodal. Algoritma ini memanfaatkan proses *birth/death* dan *split/merge* dalam aplikasinya.

Richardson dan Green [25] menjelaskan proses RJMCMC ini ke dalam enam tipe pergerakan, yaitu :

1. *updating* ω
2. *updating* θ
3. *updating* z
4. *updating* hyperparameter β
5. *split/merge* komponen-komponen *mixture*
6. *birth/death* dari komponen kosong

Proses sampai dengan langkah ke-4 tidak mengubah dimensi vektor parameter model yang terdiri dari $(\beta, \theta, k, \omega, z)$. Sedangkan langkah 5 dan 6 mengubah banyaknya komponen penyusun *mixture* k satu per satu. Pergerakan *split/merge* adalah pilihan random antara *split* ($k \rightarrow k+1$) atau *merge* ($k+1 \rightarrow k$). Adapun pergerakan dalam proses *birth/death* merupakan pilihan random antara *birth* ($k \rightarrow k+1$) atau *death* ($k+1 \rightarrow k$).

Odds Ratio

Odds ratio merupakan suatu ukuran yang digunakan untuk mengetahui tingkat resiko (kecenderungan) yaitu perbandingan antara *Odd* individu dengan kondisi variabel prediktor X pada kategori sukses dengan kategori gagal [15]. Misal X variabel kategori dengan 2 kategori, yaitu 0 dan 1, maka *odds ratio* perbandingan individu dengan kategori 1 dengan individu dengan kategori 1 yaitu

$$OR = \frac{h(t | x = 1)}{h(t | x = 0)} = \frac{h_0(t)e^\beta}{h_0(t)} = e^\beta \quad (11)$$

artinya bahwa tingkat kecepatan terjadinya *failure event* pada individu dengan kategori $x = 1$ sebesar e^β kali tingkat kecepatan terjadinya resiko peristiwa *failure event* pada individu dengan kategori $x = 0$. Sedangkan untuk variabel kontinu, nilai dari e^β mempunyai interpretasi sebagai perbandingan *odds ratio* antara individu dengan nilai x lebih besar 1 satuan dibandingkan individu yang lainnya.

Angka Lama Sekolah dan Rata-Rata Lama Sekolah

Angka lama sekolah yaitu jumlah tahun belajar penduduk usia 15 tahun ke atas yang telah diselesaikan dalam pendidikan formal (tidak termasuk tahun yang mengulang). Untuk menghitung angka lama sekolah dibutuhkan informasi berupa partisipasi sekolah, jenjang dan jenis pendidikan yang

pernah/sedang diduduki, ijazah tertinggi yang dimiliki, dan kelas tertinggi yang pernah/sedang diduduki [26]. Rumus angka lama sekolah sebagai berikut :

$$ALS = \text{Tahun Konversi} + (\text{Kelas Tertinggi} - 1) \quad (12)$$

dimana,

ALS : Angka Lama Sekolah

Tahun Konversi : Tahun Pendidikan yang Ditamatkan (SD = 6 tahun; SMP = 9 tahun; SMA = 12 tahun; D1 = 13 tahun; D2 = 12 tahun; D3 = 13 tahun; D4/S1 = 16 tahun; S2 = 18 tahun; S3 = 21 tahun)

METODOLOGI PENELITIAN

Sumber Data

Data yang dipergunakan dalam penelitian ini data sekunder bersumber dari Badan Pusat Statistik (BPS) Jawa Timur, yaitu data dari hasil survei sosial ekonomi nasional (SUSENAS) tahun 2012. Data yang digunakan sebanyak rumah tangga yang memiliki anak berumur 16-18 tahun yang berjumlah 3.794 rumah tangga.

Variabel Penelitian

Variabel respon dan variabel-variabel prediktor yang diperkirakan mempengaruhi angka lama sekolah ditunjukkan dalam tabel berikut :

Tabel 1. Daftar Variabel Yang Digunakan Dalam Penelitian

Variabel	Uraian	Skala Data	Ket.
T	Angka Lama Sekolah	Rasio	Tahun
X ₁	Jenis Kelamin	Nominal	1=Laki-Laki 2=Perempuan
X ₂	Status Perkawinan	Nominal	1=Belum Kawin 2=Kawin
X ₃	Status Bekerja	Nominal	1=Tidak Bekerja 2 = Bekerja
X ₄	Tingkat Pendidikan Kepala Rumah Tangga	Ordinal	1=TTSD 2=TSD 3=TSMP 4=TSMA 5=TPT
X ₅	Jenis Kelamin Kepala Rumah Tangga	Nominal	1=Laki-Laki 2=Perempuan
X ₆	Rata-Rata Pengeluaran Rumah Tangga per Kapita per Bulan	Rasio	Ratusan Ribu Rupiah
X ₇	Jumlah Anggota Rumah Tangga	Ratio	Jiwa
X ₈	Lokasi Tempat Tinggal	Nominal	1=Kota 2=Desa

Variabel respon yang digunakan dalam penelitian ini adalah data angka lama sekolah di Provinsi Jawa Timur tahun 2012 dari responden yang disimbolkan dengan T (yang tidak pernah menempuh sekolah, yang sedang bersekolah, dan yang pernah menempuh sekolah tapi kini tidak bersekolah), yang berumur 16-18 tahun (dalam usia sekolah menengah atas) berada dalam batas periode penelitian, dalam satuan tahun, dengan ketentuan sebagai berikut:

- Jika seorang responden pernah bersekolah ataupun masih bersekolah hingga dinyatakan lulus karena mendapatkan ijazah sekolah menengah pertama (SMP) atau yang

sederajat (MTs, SMPLB dan paket B) baik negeri maupun swasta dan dalam batas periode penelitian, maka waktu *survival* dikategorikan sebagai data *survival* tidak tersensor.

- Jika seorang responden sampai dengan batas periode penelitian, terhitung dari responden tersebut mulai bersekolah (tidak dibatasi kapanpun responden tersebut bersekolah) hingga responden tersebut mengalami hal-hal berikut:

- Masih bersekolah pada jenjang SD/sederajat atau SMP/sederajat atau melebihi batas akhir penelitian (tahun 2012),
- Jika seorang responden tidak pernah sama sekali bersekolah,
- Jika seorang responden keluar dari masa belajar (mengundurkan diri) atau mengalami *drop out* (dikeluarkan dari sekolah),

maka data *survival* dikatakan data tersensor. Variabel respon kemudian dikategorikan menjadi:

$t=0$, merupakan data tersensor jika responden tidak mengalami *failure event*. Misal adanya responden tidak pernah bersekolah sama sekali, keluar dari masa belajar (mengundurkan diri) atau mengalami *drop out* (dikeluarkan dari sekolah), atau melebihi batas akhir penelitian dan responden masih belum mendapatkan ijazah SMP/sederajat.

$t=1$, merupakan data tidak tersensor jika responden mendapatkan ijazah SMP/sederajat yang berada dalam batas periode penelitian.

Metode Analisis

Berdasarkan tujuan penelitian yang telah dijelaskan pada bab sebelumnya, maka langkah-langkah analisis dijelaskan sebagai berikut:

- Mengkaji model *survival* dengan penambahan konsep *mixture* dan pertimbangan adanya efek spasial menggunakan pendekatan bayesian.
- Menentukan model *mixture survival* spasial pada angka lama sekolah anak umur 16-18 tahun di Jawa Timur berdasarkan faktor-faktor yang mempengaruhi tinggi atau rendahnya angka lama sekolah anak umur 16-18 tahun. Langkah-langkahnya dijelaskan sebagai berikut:
 - Mengkaji karakteristik responden anak umur 16-18 tahun berdasarkan semua variabel prediktor.
 - Melakukan pengujian asumsi pemodelan *hazard* proporsional
 - Menentukan penimbang spasial
 - Mengidentifikasi dan menguji distribusi data waktu *survival* angka lama sekolah.
 - Penggunaan algoritma RJMCMC banyak komponen tidak diketahui.
 - Menentukan model dan parameter model spasial *survival* menggunakan simulasi MCMC.

HASIL DAN PEMBAHASAN

- Model *Survival* Dengan Penambahan Konsep *Mixture* dan Pertimbangan Adanya Efek Spasial Menggunakan Pendekatan Bayesian

Pendekatan lattice digunakan untuk menyusun susunan spasial dari kelompok-kelompok tersebut. Sehingga W_j terganti oleh W_j^* , hal tersebut dikarenakan W^* didefinisikan oleh daerah-daerah dengan indeks diskret dimana daerah-daerah tersebut merupakan partisi/bagian dari daerah D .

Sehingga model *hazard* proporsional pada persamaan (7) menjadi sebagai berikut :

$$h(t_{ij}; \mathbf{x}\beta) \propto h_0(t_{ij}) \exp(\mathbf{x}_{ij}^T \beta + W_j^*) \quad (13)$$

Asumsi distribusi CAR (*Conditionally Autoregressive*) untuk efek random (*frailty*) $\mathbf{W}^* | \lambda \sim \text{CAR}(\lambda)$, dimana $\lambda = 1/v$. λ juga disebut sebagai parameter penghalus yang secara implisit juga menentukan variabilitas dari efek random (*frailty*). Nilai λ akan berbeda antara daerah satu dengan daerah yang lain, hal ini bergantung pada jumlah tetangga (kedekatan dengan daerah lain) yang dimiliki oleh daerah tersebut (m_j).

Pada penelitian ini, distribusi dari angka lama sekolah anak umur 16-18 tahun mengikuti berdistribusi dengan jenis nonnegatif distribusi karena berbentuk skew kiri. Dalam pendekatan untuk setiap komponen *mixture* dari angka lama sekolah anak umur 16-18 tahun dilakukan dengan menggunakan distribusi Normal 2-parameter (μ, σ), dengan fungsi kepadatan peluang (PDF) dari distribusi Normal sebagai berikut,

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) \quad (14)$$

sedangkan fungsi distribusi kumulatifnya (CDF) sebagai berikut,

$$F(t) = \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{t-\mu}{\sqrt{2}\sigma}\right) \quad (15)$$

Pada hasil CDF pada (15) terdapat hasil erf dimana erf merupakan *eror function* dari CDF distribusi Normal dikarenakan $\frac{1}{\sqrt{2\pi}} \int_0^x \exp\left(-\frac{1}{2}x^2\right) d(x)$ memiliki hasil yang tidak *close form* sehingga terbentuklah *eror function*. Berdasarkan persamaan CDF di atas dan fungsi *survival* pada persamaan (1), maka fungsi *survival* dari distribusi Normal 2-parameter sebagai berikut,

$$S(t) = \frac{1}{2} - \frac{1}{2} \operatorname{erf}\left(\frac{t-\mu}{\sqrt{2}\sigma}\right) \quad (16)$$

fungsi *hazard* dari distribusi Normal 2-parameter sebagai berikut,

$$h(t) = \frac{\sqrt{2} \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right)}{\sqrt{\pi\sigma^2} \left(1 - \operatorname{erf}\left(\frac{t-\mu}{\sqrt{2}\sigma}\right)\right)} \quad (17)$$

Berdasarkan fungsi *hazard* pada persamaan (16) dan (4) maka dapat diperoleh persamaan berikut,

$$h(t, \mathbf{x}) = h_0(t) \exp(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p) = \frac{\sqrt{2} \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right)}{\sqrt{\pi\sigma^2} \left(1 - \operatorname{erf}\left(\frac{t-\mu}{\sqrt{2}\sigma}\right)\right)} \quad (18)$$

Pada persamaan (18), dikarenakan masih terdapat *eror function*, cukup susah untuk menentukan mana yang merupakan bagian dari $h_0(t)$ dan $\exp(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)$.

Sebelum melakukan estimasi terhadap parameter μ dan σ , maka terlebih dahulu menentukan distribusi priornya sebagai berikut,

$$\sigma \sim \text{Gamma}(p, q)$$

$$\beta_i \sim \text{Normal}(m, s)$$

dimana penentuan dimana penentuan distribusi prior tersebut dilakukan berdasarkan gabungan antara *prior conjugate* dan *prior informatif*. Pada analisis *survival* menyertakan data tersensor dalam menentukan join distribusi posteriornya, γ_{ij} menyatakan status tersensor dari responden angka lama sekolah yaitu nilai 0 untuk responden tersensor dan nilai 1 untuk responden tidak tersensor. Selanjutnya ditentukan t sebagai waktu *survival* responden (waktu yang diperlukan hingga terjadinya lulus) dan \mathbf{x} merupakan vektor *covariate* (variabel prediktor) sehingga diperoleh join distribusi posteriornya sebagai berikut,

$$P(\beta, \mathbf{W}, \sigma, \lambda | t, \mathbf{x}, \gamma) \propto L(\beta, \mathbf{W}, \sigma; t, \mathbf{x}, \gamma) p(\mathbf{W} | \lambda) p(\beta) p(\sigma) p(\lambda)$$

Pada ruas kanan terdapat lima bentuk, yaitu bentuk pertama merupakan likelihood untuk *hazard* distribusi Normal 2-parameter, bentuk kedua menyatakan join distribusi dari efek random (*frailty*), sedangkan tiga bentuk sisanya merupakan distribusi prior dari masing-masing parameter.

Pengujian dan penentuan distribusi dari data angka lama sekolah anak umur 16-18 tahun yang akan digunakan untuk model *hazard proporsional* nya, dan jika terdapat indikasi bahwa distribusi dari data angka lama sekolah anak umur 16-18 tahun terdapat *mixture* dan cukup sulit untuk menentukan komponen *mixture*nya maka akan digunakan algoritma *Reversible Jump Markov Chain Monte Carlo* (RJMCMC). Sedangkan untuk menentukan parameter pada model dengan menggunakan MCMC dengan *Gibbs Sampling*.

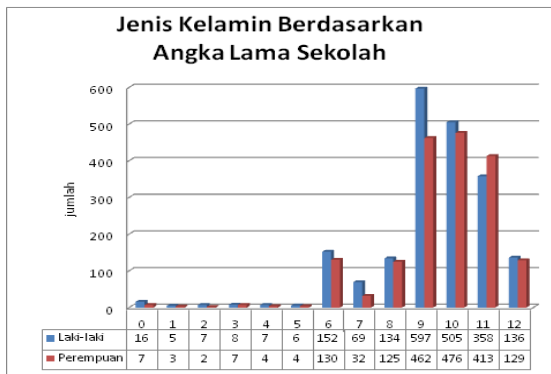
2. Model *Mixture Survival* Spasial Pada Angka Lama Sekolah Anak Umur 16-18 Tahun di Jawa Timur Berdasarkan Variabel-Variabel Yang Mempengaruhi.

Karakteristik responden anak umur 16-18 tahun berdasarkan variabel-variabel prediktornya dapat dilihat pada tabel dan gambar berikut.

Tabel 2. Deskriptif Responden Anak Berumur 16-18 Tahun

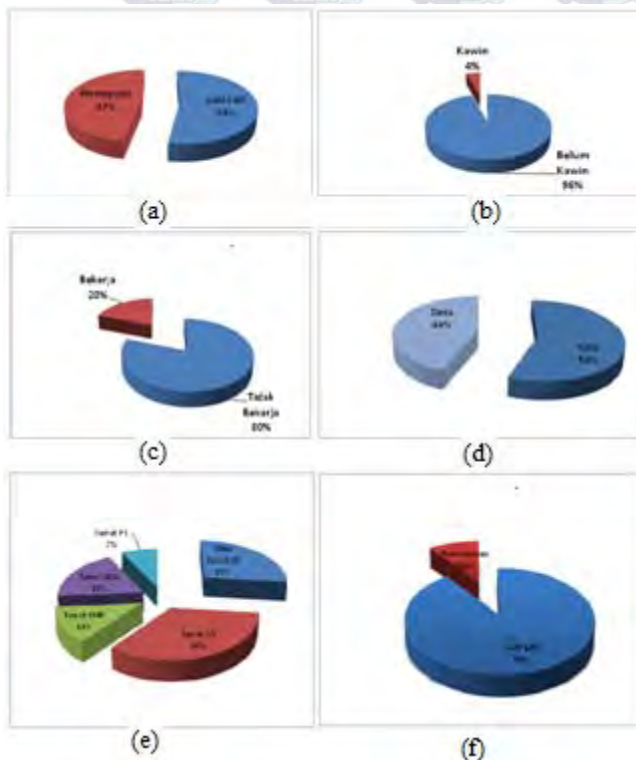
Variabel	Minimum	Maksimum	Mean	Std. Deviation
Angka Lama Sekolah Anak Umur 16-18 Tahun	0,00	12,00	9,39	1,84
Jumlah Anggota Rumah Tangga	1,00	11,00	4,41	1,34
Rata-Rata Pengeluaran Rumah Tangga	1,21	131,13	5,17	4,81

Pada Tabel 2, waktu *survival* responden atau angka lama sekolah anak umur 16-18 tahun paling rendah 0 tahun yang artinya belum pernah menyentuh pendidikan formal. Paling tinggi yaitu 12 tahun yang artinya sudah mendapatkan ijazah pada jenjang SMA/ sederajat, sedangkan mean atau rata-rata nya 9,39 yang artinya sebagian besar anak yang berumur 16-18 tahun memiliki rata-rata lama sekolah di atas 9 tahun atau sudah lulus pada jenjang SMP/ sederajat. Untuk variabel jumlah anggota keluarga, paling sedikit ada rumah tangga yang memiliki 1 anggota keluarga dan yang paling banyak hingga 11 anggota keluarga di dalamnya. Sedangkan untuk variabel rata-rata pengeluaran rumah tangga per kapita per bulan, pengeluaran terendah yaitu Rp.120.955,94 dan pengeluaran tertinggi hingga Rp.13.113.030,12, sedangkan untuk rata-rata dari rata-rata pengeluaran rumah tangga yaitu Rp.517.887,26.



Gambar 1. Diagram Batang Variabel Jenis Kelamin Berdasarkan Angka Lama Sekolah Anak Umur 16-18 Tahun

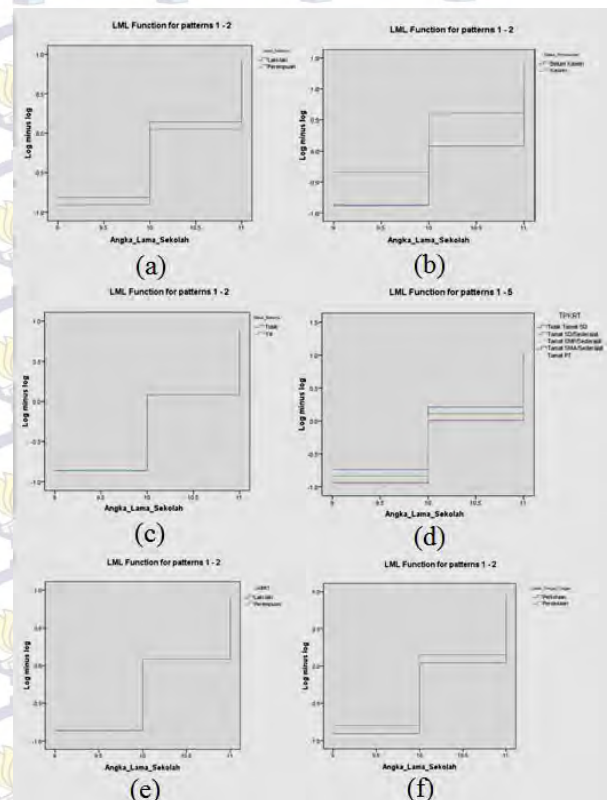
Perbedaan jumlah responden laki-laki dan perempuan disetiap jenjang angka lama sekolah dapat dilihat pada Gambar 1, bahwa disetiap jenjang angka lama sekolah responden laki-laki (warna biru) lebih banyak dari responden perempuan (warna merah), kecuali pada angka lama sekolah 11 dimana jumlah responden perempuan lebih banyak.



Gambar 2. Pie Chart Variabel Prediktor Kategorik (a). Jenis Kelamin Responden, (b). Status Perkawinan Responden, (c). Status Bekerja Responden, (d). Lokasi Tempat Tinggal Responden, (e). Tingkat Pendidikan Terakhir Kepala Rumah Tangga, dan (f). Jenis Kelamin Kepala Rumah Tangga

Pada Gambar 2(a) memberikan informasi bahwa jumlah responden anak laki-laki umur 16-18 tahun (warna biru) lebih banyak dari responden perempuan (warna merah), hal tersebut dapat dilihat jumlah responden laki-laki mencapai 53% dari jumlah responden anak umur 16-18 tahun secara keseluruhan sedangkan untuk responden perempuan hanya 47%. Berdasarkan Gambar 2(b), terdapat 4% responden yang sudah kawin (warna merah) dan sebagian besar dengan 96% belum kawin (warna biru). Gambar 2(c) memberikan informasi bahwa terdapat 20% responden yang sudah bekerja (warna merah) pada umur 16-18 tahun sedangkan sisanya 80% tidak bekerja (warna biru).

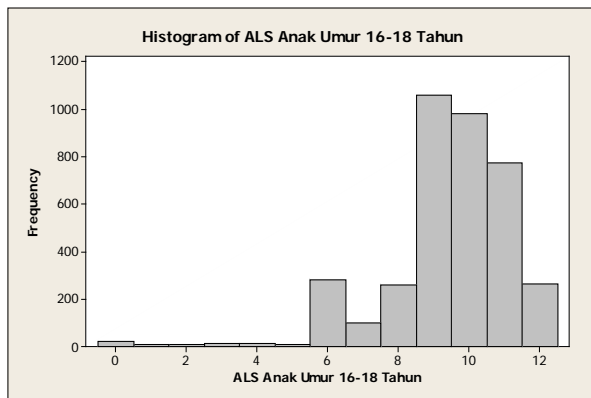
Pada Gambar 2(d), lokasi tempat tinggal responden, 56% dari total responden bertempat tinggal di perkotaan (warna biru), sedangkan sisanya 44% dari total responden bertempat tinggal di perdesaan (warna abu-abu). Gambar 2(e) menjelaskan bahwa 27 % tidak tamat SD/ sederajat atau dapat dikatakan tidak mendapat pendidikan formal (warna biru), 34% hanya memiliki ijazah sampai jenjang SD/ sederajat (warna merah), 13% memiliki ijazah sampai jenjang SMP/ sederajat (warna hijau), 19% memiliki ijazah sampai jenjang SMA/ sederajat (warna ungu) dan sisanya 7% memiliki ijazah sampai perguruan tinggi (warna biru muda). Sedangkan Gambar 2(f) menjelaskan kepala rumah tangga berjenis kelamin laki-laki (warna biru) lebih banyak dari kepala rumah tangga berjenis kelamin perempuan (warna merah), yaitu 90% (3410 kepala rumah tangga) yang berjenis kelamin laki-laki dan 10% (384 kepala rumah tangga) yang berjenis kelamin perempuan.



Gambar 3. Plot $\ln[-\ln S(t)]$ Uji Proporsional Hazard (a). Jenis Kelamin Responden, (b). Status Perkawinan Responden, (c). Status Bekerja Responden, (d). Tingkat Pendidikan Terakhir Kepala Rumah Tangga, (e). Jenis Kelamin Kepala Rumah Tangga, dan (f). Lokasi Tempat Tinggal Responden.

Dari Gambar 3. dapat dilihat bahwa garis antar kategori sejajar sehingga bisa dikatakan asumsi *hazard* proporsional sudah terpenuhi untuk keenam variabel kategorik, artinya variabel prediktor telah independen terhadap waktu *survival* serta hubungan antara kumulatif *hazard* sudah proporsional/ konstan setiap waktu.

Penentuan matriks penimbang spasial digunakan penimbang *Queen Contiguity* (persinggungan sisi dan sudut) dan kabupaten yang saling bertetangga disebut sebagai *adjacent* sehingga ketetanggaan dari masing-masing kabupaten di provinsi Jawa Timur disusun dalam matriks *adjacent* (matriks penimbang terdapat pada syntax Lampiran 2). Matriks penimbang spasial yang diperoleh selanjutnya merupakan salah satu parameter dalam prior CAR yang menjadi distribusi dari efek random (*frailty*) model *survival*.



Gambar 4. Histogram Angka Lama Sekolah Anak Umur 16-18 Tahun

Tabel 3. Uji Distribusi Waktu *Survival*

Distribusi	Statistik Uji (A_n^2)	Nilai Kritis ($a_{n,1-\alpha}$)	Keputusan
Log Normal	281,78	2,5018	Tolak H_0
Normal	146,06	2,5018	Tolak H_0
Weibull	147,00	2,5018	Tolak H_0
Log Logistik	193,50	2,5018	Tolak H_0

Jika dilihat berdasarkan pada Tabel 3. bahwa tidak ada satupun distribusi yang sesuai untuk distribusi angka lama sekolah anak umur 16-18 tahun, hal tersebut dikarenakan adanya multimodal sehingga saat pengujian distribusi waktu *survival* tidak ada yang signifikan. Hal tersebut didukung pada Gambar 4., yang terdapat lebih dari 1 puncak pada histogram.

Penentuan banyak komponen yang dibentuk oleh model dengan menggunakan algoritma RJMCMC (Syntax algoritma RJMCMC pada Lampiran 1), dengan menggunakan iterasi sebanyak 100000 didapatkan banyak komponen untuk distribusi *mixture* pada data *survival* angka lama sekolah anak umur 16-18 tahun yaitu sebanyak 6 komponen. Nilai komponen awal setiap *mixture*, yaitu sebesar 0,86021, 0,05878, 0,05303, 0,01396, 0,00557, dan 0,00845. Nilai komponen awal ini digunakan untuk menentukan data yang sesuai dengan komponennya.

Jumlah komponen *mixture* yang telah diketahui digunakan untuk membentuk model *mixture survival* spasial dengan variabel-variabel yang mempengaruhi. Digunakan software WinBUGS untuk membantu dalam perhitungan serta penentuan nilai parameter dan variabel-variabel yang mempengaruhi (Syntax pada Lampiran 2).

Hasil estimasi parameter (*posterior summaries*) model *mixture survival* spasial bedistribusi Normal disajikan pada Lampiran 3. Variabel-variabel di atas dianggap signifikan mempengaruhi laju putus sekolah (berhenti sekolah) responden anak umur 16-18 tahun jika nilai pada selang 2,5% hingga 97,5% tidak memuat nilai 0 (pada Lampiran 3 yang berwarna kuning tidak berpengaruh secara signifikan). Lampiran 3 menunjukkan bahwa tidak semua variabel berpengaruh signifikan terhadap laju putus sekolah (berhenti sekolah) responden anak umur 16-18 tahun dan terdapat komponen *mixture* yang dipengaruhi oleh beberapa variabel saja. Kolom *node* merupakan kovariat (variabel-variabel) yang diduga berpengaruh terhadap laju putus sekolah (berhenti sekolah) responden anak umur 16-18 tahun, kolom *mean* menunjukkan besarnya nilai parameter model, sedangkan dua kolom berikutnya SD dan MC error menunjukkan besar varian besarnya nilai estimasi pada selang kepercayaan 95%.

Pada komponen *mixture* pertama dengan besar proporsi 0,00869, responden yang masuk yaitu responden dengan angka lama sekolah 0, 1 dan 2. Variabel-variabel yang secara signifikan mempengaruhi model meliputi status perkawinan responden (X2), status bekerja responden (X3), jenis kelamin kepala rumah tangga (X5), dan lokasi tempat tinggal responden (X8). Untuk komponen *mixture* kedua dengan besar proporsi 0,00767, responden yang masuk yaitu responden dengan angka lama sekolah 1, 2, 3, 4, dan 5, serta aktor-variabel yang secara signifikan mempengaruhi model meliputi semua variabel kecuali variabel dummy kategori 1 pada variabel tingkat pendidikan tertinggi kepala keluarga. Untuk komponen *mixture* ketiga dengan besar proporsi 0,0642, responden yang masuk yaitu responden dengan angka lama sekolah 4, 5, 6, 7, dan 8. Variabel-variabel yang secara signifikan mempengaruhi model meliputi semua variabel kecuali variabel dummy kategori 1 dan kategori 4 pada variabel tingkat pendidikan tertinggi kepala keluarga. Untuk komponen *mixture* keempat dengan besar proporsi 0,09174, responden yang masuk yaitu responden dengan angka lama sekolah 6, 7, 8, 9, dan 10, dimana semua variabel secara signifikan mempengaruhi model pada *mixture* keempat. Untuk *mixture* kelima dengan besar proporsi 0,7984, responden yang masuk yaitu responden dengan angka lama sekolah 8, 9, 10, 11, dan 12, dimana sama hal nya seperti *mixture* keempat dimana semua variabel secara signifikan mempengaruhi model pada *mixture* kelima. Sedangkan untuk *mixture* keenam dengan besar proporsi 0,02929, responden yang masuk yaitu responden dengan angka lama sekolah 10, 11, dan 12. Variabel-variabel yang secara signifikan mempengaruhi model pada *mixture* keenam meliputi jenis kelamin responden (X1), status perkawinan responden (X2), status bekerja responden (X3), jenis kelamin kepala rumah tangga (X5), jumlah anggota rumah tangga (X7), dan lokasi tempat tinggal responden (X8).

Untuk penginterpretasian estimasi parameter model digunakan nilai odds ratio dengan perhitungan \exp dari nilai parameter ($\exp(b)$). Sebagai contoh pada *mixture* ke-4, dimana pada *mixture* ke-4 semua parameter signifikan berpengaruh terhadap angka lama sekolah. Pada *mixture* ke-4 dengan proporsi 0,092, varian (ptau) 0,27 dan respon yang termasuk dalam *mixture* ke-4 respon dengan angka lama sekolah 6, 7, 8, 9, dan 10.

Pada variabel jenis kelamin responden, tingkat kecenderungan responden dengan jenis kelamin laki-laki untuk memperoleh kesempatan sekolah lebih lama sebesar $\exp(b1[4]) = 1,1$ kali daripada responden dengan jenis kelamin perempuan, artinya jika ada responden perempuan dengan angka lama sekolah 8 tahun maka responden laki-laki memiliki angka lama sekolah 9 tahun.

Untuk variabel status perkawinan responden, tingkat kecenderungan responden yang berstatus belum kawin untuk memperoleh kesempatan sekolah lebih lama sebesar $\exp(b2[4]) = 1,41$ kali daripada responden yang sudah kawin, artinya jika ada responden dengan status kawin putus sekolah pada angka lama sekolah 7 tahun maka responden yang belum kawin akan memiliki angka lama sekolah 10 tahun. Hal tersebut dikarenakan jika seorang anak kawin pada jenjang pendidikan SD, SMP, atau SMA, anak tersebut dikeluarkan dari sekolah/*drop out*, maka responden tersebut sulit mendapatkan kesempatan sekolah lebih lama.

Untuk variabel status bekerja responden, tingkat kecenderungan responden yang berstatus belum bekerja untuk memperoleh kesempatan sekolah lebih lama sebesar $\exp(b3[4]) = 1,06$ daripada responden yang bekerja, artinya jika ada responden dengan status bekerja pada usia sekolah dengan angka lama sekolah 9 tahun maka responden yang belum

bekerja akan memiliki angka lama sekolah 10 tahun. Hal tersebut dikarenakan jika seorang anak bekerja pada usia sekolah maka anak tersebut akan lebih cenderung untuk memprioritaskan pekerjaan daripada sekolah.

Pada variabel tingkat pendidikan terakhir kepala rumah tangga dari responden, kepala rumah tangga dengan status pendidikan tertinggi tidak tamat SD/ sederajat lebih cenderung memiliki anak umur 16-18 tahun memperoleh kesempatan bersekolah lebih lama sebesar $\exp(b4_1[4]) = 3,65$ kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain, untuk kepala rumah tangga dengan status pendidikan tertinggi tamat SD/ sederajat lebih cenderung memiliki anak umur 16-18 tahun memperoleh kesempatan bersekolah lebih lama sebesar $\exp(b4_2[4]) = 3,53$ kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain, untuk kepala rumah tangga dengan status pendidikan tertinggi tamat SMP/ sederajat lebih cenderung memiliki anak umur 16-18 tahun memperoleh kesempatan bersekolah lebih lama sebesar $\exp(b4_3[4]) = 3,96$ kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain, sedangkan untuk kepala rumah tangga dengan status pendidikan tertinggi tamat SMA/ sederajat lebih cenderung memiliki anak umur 16-18 tahun memperoleh kesempatan bersekolah lebih lama sebesar $\exp(b4_4[4]) = 3,68$ kali daripada kepala rumah tangga dengan tingkat pendidikan yang lain.

Pada variabel jenis kelamin kepala rumah tangga dari responden, tingkat kecenderungan kepala rumah tangga dengan jenis kelamin laki-laki memiliki anak umur 16-18 tahun untuk memperoleh kesempatan bersekolah lebih lama sebesar $\exp(b5[4]) = 1,16$ kali daripada responden dengan kepala rumah tangga berjenis kelamin perempuan, artinya jika ada responden dengan kepala rumah tangga perempuan memiliki angka lama sekolah 7 tahun, maka responden lainnya dengan kepala rumah tangga laki-laki akan memiliki angka lama sekolah minimal sebesar 8 tahun. Responden dengan kepala rumah tangga perempuan akan lebih cenderung untuk bekerja juga membantu keuangan rumah tangga sehingga untuk keperluan sekolah dikesampingkan.

Pada variabel pengeluaran rumah tangga per kapita per bulan, setiap penambahan 1 satuan pengeluaran per kapita per bulan rumah tangga lebih cenderung seorang responden memperoleh kesempatan bersekolah lebih lama sebesar $\exp(b6[4]) = 1,07$ kali atau dalam interpretasi lain, setiap penambahan 1 satuan pengeluaran per kapita per bulan rumah tangga lebih cenderung membuat keinginan anak umur 16-18 tahun pada rumah tangga tersebut sebesar 1,07 kali untuk tetap bersekolah. Untuk variabel jumlah anggota rumah tangga, setiap penambahan 1 satuan anggota rumah tangga lebih cenderung membuat keinginan anak umur 16-18 tahun pada rumah tangga tersebut sebesar $\exp(b7[4]) = 1,02$ kali untuk mendapatkan kesempatan bersekolah lebih sedikit atau tidak melanjutkan pendidikan untuk jenjang pendidikan yang lebih tinggi dari SMP/ sederajat.

Sedangkan untuk variabel lokasi tempat tinggal responden, tingkat kecenderungan responden dengan lokasi tempat tinggal di perkotaan untuk memperoleh kesempatan sekolah lebih lama sebesar $\exp(b8[4]) = 1,07$ kali daripada responden dengan lokasi tempat tinggal di perdesaan, artinya jika terdapat responden dengan lokasi tempat tinggal di perdesaan memiliki angka lama sekolah 8 tahun maka anak yang tinggal di daerah perkotaan akan memiliki angka lama sekolah minimal 9 tahun.

Pada Lampiran 4 dapat dilihat bahwa baris yang berwarna kuning adalah daerah-daerah dengan efek spasial yang tidak signifikan. Signifikan atau tidaknya dapat dilihat

dari nilai pada selang 2.5% hingga 97.5% tidak memuat nilai 0.

Daerah-daerah dengan efek spasial yang secara signifikan mempengaruhi angka lama sekolah anak umur 16-18 tahun artinya bahwa pengaruh spasial atau pengaruh kedaerahan, dimana dalam penelitian ini pengaruh spasial yang dianggap random (menjadi variabel random W_j^*), memiliki efek yang berbeda disetiap daerah dan mempengaruhi laju putus sekolah (berhenti sekolah sebelum SMP) disetiap daerah secara berbeda. Hal tersebut dikarenakan setiap daerah mungkin saja memiliki kebijakan yang berbeda atau keadaan yang berbeda yang diciptakan oleh masyarakat antara satu daerah dengan daerah lain sehingga mempengaruhi laju putus sekolah (berhenti sekolah) di daerah tersebut yang berimbas bahwa rata-rata lama sekolah berdasarkan angka lama sekolah anak umur 16-18 tahun di daerah tersebut berbeda juga.

Jadi model *mixture survival* spasial (*hazard rate*) yang terbentuk yaitu

$$h(t, \mathbf{x}) = 0,087(h_1(t) \exp(0,98x_2 + 1,33x_3 + 0,41x_5 + 0,88x_8 + W_j^*)) \\ + 0,008(h_2(t) \exp(0,53x_1 + 0,38x_2 - 1,03x_3 + 0,62x_{4_2} \\ - 1,33x_{4_3} - 1,77x_{4_4} - 0,8x_5 + 0,16x_6 + 0,8x_7 + 0,55x_8 + W_j^*)) \\ + 0,06(h_3(t) \exp(0,54x_1 + 0,25x_2 + 0,17x_3 + 0,27x_{4_2} - 1,02x_{4_3} \\ - 0,22x_5 + 0,57x_6 + 0,37x_7 + 0,02x_8 + W_j^*)) \\ + 0,092(h_4(t) \exp(0,1x_1 + 0,34x_2 + 0,06x_3 + 1,29x_{4_1} + 1,26x_{4_2} \\ + 1,38x_{4_3} + 1,3x_{4_4} + 0,15x_5 + 0,07x_6 + 0,02x_7 + 0,07x_8 + W_j^*)) \\ + 0,798(h_5(t) \exp(0,14x_1 + 0,23x_2 + 0,17x_3 + 0,6x_{4_1} + 0,62x_{4_2} \\ + 0,67x_{4_3} + 0,69x_{4_4} + 0,26x_5 + 0,013x_6 + 0,072x_7 + 0,2x_8 + W_j^*)) \\ + 0,029(h_6(t) \exp(0,122x_1 + 0,73x_2 + 0,26x_3 + 0,15x_{4_2} + 0,2x_5 \\ + 0,05x_7 + 0,52x_8 + W_j^*))$$

dimana $h_1(t)$ hingga $h_6(t)$ adalah fungsi baseline hazard dari setiap komponen *mixture* dan W_j^* adalah efek spasial dari daerah ke- j .

PENUTUP

Kesimpulan

Berdasarkan hasil yang diperoleh dapat disimpulkan bahwa

1. Model umum *mixture survival* spasial dimana pendekatan setiap komponen *mixture* dengan distribusi normal didapatkan 6 komponen *mixture* dan efek spasial W_j^* didapatkan sebagai berikut,

$$h(t, \mathbf{x}) = \pi_1(h_1(t) \exp(\beta_1^T x_{p1} + W_j^*)) + \pi_2(h_2(t) \exp(\beta_2^T x_{p2} + W_j^*)) \\ + \pi_3(h_3(t) \exp(\beta_3^T x_{p3} + W_j^*)) + \pi_4(h_4(t) \exp(\beta_4^T x_{p4} + W_j^*)) \\ + \pi_5(h_5(t) \exp(\beta_5^T x_{p5} + W_j^*)) + \pi_6(h_6(t) \exp(\beta_6^T x_{p6} + W_j^*))$$

2. Variabel-variabel yang berpengaruh secara signifikan berbeda pada 6 komponen *mixture*. Serta terdapat 23 kabupaten/kota memiliki efek spasial yang secara signifikan mempengaruhi angka lama sekolah anak umur 16-18 tahun.

Saran

Pada penelitian selanjutnya diharapkan agar data untuk model *mixture survival* spasial menggunakan data panel agar selain ada efek spasial terdapat juga efek waktu yang mempengaruhi variabel respon. Perlu juga dilakukan pengembangan dalam bidang spasial, agar yang digunakan tidak hanya efek spasial namun dependensi spasial serta pengembangan dalam Reversible Jump Markov Chain Monte Carlo untuk distribusi selain distribusi Normal.

DAFTAR PUSTAKA

- [1] Collet, D. (1994). *Modelling Survival Data in Medical Research*. London: Chapman and Hall.
- [2] Marin, J. M., Mengersen, K. & Robert, C. P., (2001). Bayesian Modelling and Inderence on *Mixture of Distribution*. *Handbook of Statistics*, 25(50).
- [3] Darmofal, D. (2008). *Bayesian Spatial Survival Models for Political Event Processes*. Columbia: Departement of Political Science, University of South Carolina 350 Gambrel Hall.
- [4] Banerjee, S., Wall, M. M., & Carlin, B. P. (2003). Frailty Modeling for Spatially Correlated *Survival Data*, With Application to Infant Mortality in Minnesota. *Biostatistics* , 123-142.
- [5] Aksioma, D. F. (2011). *Model Spasial Survival dengan Pendekatan Bayesian (Studi Kasus pada Kejadian HIV/AIDS di Provinsi Jawa Timur)*. Surabaya: Institut Teknologi Sepuluh Nopember.
- [6] Hasyim, M. (2012). *Model Mixture Survival Spasial dengan Frailty Berdistribusi Conditionally Autoregressive (CAR) pada Kasus Kejadian Demam Berdarah Dengue (DBD) di Kabupaten Pamekasan*. Surabaya: Institut Teknologi Sepuluh Nopember.
- [7] BPS. (2012). *Statistik Indonesia 2012*. Jakarta: BPS.
- [8] Santoso, B. (2009). *Pendekatan Spline Multivariable Dan MARS Untuk Pemodelan Lama Sekolah Pada Penduduk Usia Sekolah di Provinsi Papua*. Tesis. Surabaya: Institut Teknologi Sepuluh Nopember.
- [9] Sulistiyawati, D. (2009). *Model Mixture Survival Pada Kasus Lama Sekolah di Kabupaten Boalemo Provinsi Gorontalo*. Tesis. Surabaya: Institut Teknologi Sepuluh Nopember.
- [10] Kleinbaum, D. G., & Klein, M. (2005). *Survival Analysis: A Self Learning* (3rd ed.). New York: Springer.
- [11] Zang. (2008). *Survival Analysis*. California: Wadsworth.
- [12] Vittinghoff, E., Glidden, D. V., Shiboski, S. C., & McCulloch, C. E. (2005). *Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models*. New York: Springer.
- [13] Collet, D. (1994). *Modelling Survival Data in Medical Research*. London: Chapman and Hall.
- [14] Collet, D. (2003). *Modelling Survival Data in Medical Research 2nd ed*. London: Chapman and Hall.
- [15] Hosmer Jr., D., & Lemeshow, S., (1999). *Applied Survival Analysis Regression Modelling of Time to Event Data*. New York: John Wiley and Sons. Inc.
- [16] Wall, M. M. (2004). A Close Look at The Spatial Structure Implied by the CAR and SAR Models. *Journal of Statistical Planning and Inference*, 121 , 311-324.
- [17] Lesage, J. P. (1999). *The Theory and Practice of Spatial Econometrics*. Toledo: Department of Economics, University of Toledo.

- [18] McLachlan, G., & Basford, K. E. (1988). *Mixture Models Inference and Application to Clustering*. New York: Marcel Dekker.
- [19] Iriawan, N. (2001). *Studi Tentang Bayesian Mixture Normal dengan Menggunakan Metode MCMC*. Surabaya: Lemlit ITS.
- [20] Boldstad, W.M., (2007). *Introduction to Bayesian Statistics. 2nd Edition*. New Jersey: Wiley.
- [21] Box, G. E. & Tiao, G. C., (1973). *Bayesian Inference in Statistical Alaysis*. London: Addison-Wisley.
- [22] Carlin, B. P., & Louis, T. A. (2000). *Bayes And Empirical Bayes Methods For Data Analysis (2ed)*. Boca Raton, FL: Chapman and Hall/CRC Press.
- [23] Ntzoufras, I. (2009). *Bayesian Modelling Using WinBUGS*. New York: John Wiley and Sons, Inc.
- [24] Green, P. J., 1995. Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination. *Biometrika*, 82(4), pp. 711-732.
- [25] Richardson, S., & Green, P. J. (1997). On Bayesian Analysis of *Mixture* with an Unknown Number of Component. *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 59, No. 4 , 731-792.
- [26] BPS. (2014). *sirusa.bps.go.id*. Retrieved October 6, 2014, from *sirusa.bps.go.id*: <http://sirusa.bps.go.id>

LAMPIRAN

Lampiran 1

Syntax Algoritma RJMCMC Distribusi Normal

```
model {  
  for (i in 1:n) {  
    T[i] ~ dnorm(psi[i], tau)I(t.cen[i],)  
  }  
  psi[1:n] <- jump.lin.pred(X[1:n, 1:Q], k, beta.prec)  
  id <- jump.model.id(psi[1:n])  
  beta.prec <- tau / lambda  
  tau ~ dgamma(a, b)  
  k ~ dbin(0.5, Q)}  
  
list(  
  n = 3794, Q = 11,  
  lambda = 10000,  
  a = 0.001, b = 0.001,  
  T = c(0,0,...,12,12), t.cen=c(0,0,...,1,1),  
  X = structure(  
    .Data = c(  
      2,1,1,1,0,0,0,1,1,1.89602,2,  
      1,1,2,1,0,0,0,1,6,2.51399,2,  
      :  
      :  
      1,1,1,0,1,0,0,1,5,4.97648,1,  
      1,1,2,0,1,0,0,1,4,3.19188,1  
    ),  
    .Dim = c(3794, 11)))  
  
list(tau = 1)
```


Syntax *Mixture Survival* Spasial 6 Komponen

[illegible]

Hasil Estimasi Parameter Model *Mixture Survival* Spasial

Node	Mean	Sd	MC Error	2.5%	Median	97.5%
Phi[1]	0.0087	0.0014	0.0332	0.0059	0.0086	0.0118
Phi[2]	0,0077	0,0013	0,0279	0,0052	0,00758	0,0106
Phi[3]	0,0642	0,0040	0,0913	0,0566	0,0641	0,072
Phi[4]	0,0917	0,0046	0,0001	0,0829	0,0916	0,101
Phi[5]	0,7984	0,0065	0,133	0,7856	0,7986	0,811
Phi[6]	0,0293	0,0027	0,0561	0,0241	0,02919	0,035
b1[1]	0,0271	0,0893	0,0127	-0,1669	0,0249	0,196
b2[1]	0,9876	0,1508	0,0221	0,7003	0,9863	1,288
b3[1]	1,3280	0,1621	0,0242	1,0910	1,3200	1,599
b4_1[1]	-0,087	0,1801	0,0263	-0,5033	-0,0793	0,202
b4_2[1]	0,078	0,3273	0,0438	-0,6468	0,0727	0,758
b4_3[1]	-0,180	0,2108	0,0302	-0,5683	-0,1531	0,172
b4_4[1]	-0,087	0,3676	0,0499	-0,7417	-0,1116	0,796
b5[1]	0,4103	0,0929	0,0131	0,2095	0,4080	0,564
b6[1]	0,0037	0,0216	0,0029	-0,0382	0,8450	0,057
b7[1]	0,0256	0,0202	0,0028	-0,0748	-0,0232	0,004
b8[1]	-0,877	0,0850	0,0123	-1,0150	-0,8855	-0,698
b1[2]	0,5267	0,0262	0,0039	0,4790	0,5311	0,561
b2[2]	0,3795	0,0173	0,0026	0,3456	0,3817	0,408
b3[2]	-1,028	0,0147	0,0021	-1,0510	-1,0320	-0,99
b4_1[2]	0,0243	0,0720	0,0107	-0,0926	0,0442	0,12
b4_2[2]	0,6229	0,0595	0,0089	0,5324	0,6130	0,713
b4_3[2]	-1,329	0,2244	0,0256	-1,8250	-1,3200	0,8784
b4_4[2]	-1,766	0,2069	0,0258	-2,2410	-1,7680	1,4300
b5[2]	-0,804	0,0220	0,0032	-0,8425	-0,8030	0,7690
b6[2]	0,1589	0,0057	0,8440	0,1485	0,1600	0,1685
b7[2]	0,7912	0,0025	0,3610	0,7861	0,7922	0,7945
b8[2]	0,5496	0,0095	0,0014	0,5362	0,5468	0,5681
b1[3]	0,5389	0,0116	0,0017	0,5177	0,5380	0,5625
b2[3]	0,2462	0,0104	0,0015	0,2274	0,2447	0,2639
b3[3]	0,1697	0,0145	0,0021	0,1403	0,1742	0,1933
b4_1[3]	0,0103	0,0193	0,0028	-0,0166	0,0045	0,0466
b4_2[3]	0,2694	0,0204	0,0030	0,2300	0,2701	0,3010
b4_3[3]	-1,024	0,1288	0,0181	-1,266	-1,0210	0,7419
b4_4[3]	0,1204	0,3459	0,0514	-0,579	0,2927	0,5987
b5[3]	-0,224	0,0167	0,0025	-0,251	-0,2278	0,1948
b6[3]	0,5691	0,0027	0,4010	0,5652	0,5688	0,5729
b7[3]	0,3746	0,0036	0,5390	0,369	0,3743	0,3811
b8[3]	0,0199	0,0044	0,6320	0,0107	0,0202	0,0270
b1[4]	0,0972	0,0389	0,0057	0,0265	0,0956	0,1833
b2[4]	0,3460	0,0511	0,0075	0,2371	0,3584	0,4167
b3[4]	0,0562	0,0372	0,005	0,0014	0,0483	0,1394
b4_1[4]	1,294	0,1004	0,015	1,1360	1,2720	1,4920

Node	Mean	Sd	Mc Error	2.5%	Median	97.5%
b4_2[4]	1,262	0,1105	0,016	1,0890	1,2320	1,5150
b4_3[4]	1,375	0,1336	0,019	1,1490	1,3830	1,6170
b4_4[4]	1,302	0,1066	0,0157	1,103	1,298	1,514
b5[4]	0,149	0,0814	0,012	0,0139	0,1631	0,2691
b6[4]	0,07	0,0024	0,35	0,0652	0,0711	0,0736
b7[4]	0,023	0,012	0,002	0,0018	0,0233	0,0453
b8[4]	0,07	0,0331	0,003	0,0107	0,0671	0,1427
b1[5]	0,142	0,0111	0,001	0,1194	0,1418	0,1627
b2[5]	0,235	0,0151	0,002	0,203	0,237	0,2607
b3[5]	0,169	0,0097	0,001	0,1528	0,1687	0,1882
b4_1[5]	0,6	0,0316	0,004	0,5353	0,602	0,6664
b4_2[5]	0,62	0,0325	0,005	0,5482	0,6193	0,691
b4_3[5]	0,67	0,0315	0,004	0,6084	0,6691	0,7387
b4_4[5]	0,69	0,03	0,004	0,6328	0,69	0,7594
b5[5]	0,262	0,0141	0,002	0,2347	0,2604	0,2922
b6[5]	0,01	0,0006	0,039	0,0121	0,0133	0,0144
b7[5]	0,07	0,0038	0,508	0,0647	0,0717	0,0799
b8[5]	0,2	0,0107	0,001	0,1821	0,2045	0,2215
b1[6]	0,12	0,0281	0,004	0,0583	0,124	0,1707
b2[6]	0,73	0,0918	0,013	0,5384	0,7199	0,912
b3[6]	0,2573	0,03	0,0035	0,2015	0,2576	0,311
b4_1[6]	-0,05	0,0818	0,01	-0,2002	-0,0462	0,1035
b4_2[6]	0,153	0,0635	0,008	0,0397	0,1538	0,2768
b4_3[6]	0,09	0,064	0,007	-0,0201	0,095	0,2263
b4_4[6]	0,05	0,0554	0,007	-0,0491	0,0492	0,1588
b5[6]	0,19	0,0389	0,005	0,1236	0,1953	0,2728
b6[6]	0,005	0,0027	0,271	-0,816	0,005	0,0098
b7[6]	0,054	0,0124	0,001	0,0262	0,0547	0,0779
b8[6]	0,52	0,062	0,009	0,3849	0,5171	0,6355
ptau[1]	3,66	4,141	0,443	0,1324	2,147	15,8
ptau[2]	0,196	0,406	0,04869	0,001	0,04954	1,423
ptau[3]	0,012	0,0063	0,0007	0,004	0,01007	0,03
ptau[4]	0,27	0,04	0,003	0,198	0,2693	0,36
ptau[5]	0,14	0,004	0,161	0,1317	0,1394	0,15
ptau[6]	0,37	0,05	0,003	0,2781	0,3719	0,5
sigma	0,36	0,04	0,00135	0,2837	0,3522	0,45

Lampiran 4.

Node	Mean	Sd	Mc Error	2.5%	Median	97.5%
W[1]	-0,272	0,045	0,006	-0,378	-0,267	-0,191
W[2]	-0,348	0,038	0,006	-0,413	-0,357	-0,285
W[3]	0,002	0,033	0,003	-0,060	0,002	0,067
W[4]	-0,076	0,023	0,003	-0,119	-0,076	-0,027
W[5]	-0,022	0,031	0,004	-0,079	-0,023	0,040
W[6]	-0,009	0,024	0,003	-0,057	-0,009	0,034
W[7]	-0,030	0,028	0,003	-0,078	-0,032	0,027
W[8]	0,058	0,031	0,003	0,987	0,058	0,121
W[9]	0,066	0,024	0,003	0,023	0,065	0,112
W[10]	-0,014	0,034	0,005	-0,070	-0,021	0,063
W[11]	0,046	0,034	0,004	-0,016	0,044	0,121
W[12]	0,086	0,039	0,005	0,017	0,086	0,165
W[13]	0,005	0,023	0,003	-0,042	0,005	0,052
W[14]	-0,055	0,037	0,005	-0,110	-0,069	0,010
W[15]	0,082	0,024	0,002	0,033	0,082	0,129
W[16]	-0,010	0,022	0,002	-0,052	-0,010	0,032
W[17]	0,091	0,024	0,002	0,042	0,090	0,137
W[18]	-0,043	0,025	0,004	-0,085	-0,045	-0,003
W[19]	0,011	0,027	0,002	-0,041	0,008	0,070
W[20]	0,041	0,038	0,002	-0,039	0,043	0,112
W[21]	0,039	0,030	0,003	-0,019	0,038	0,096
W[22]	-0,217	0,017	0,002	-0,249	-0,214	-0,192
W[23]	-0,014	0,026	0,002	-0,064	-0,013	0,039
W[24]	-0,171	0,028	0,004	-0,211	-0,178	-0,115
W[25]	0,046	0,023	0,002	0,001	0,046	0,091
W[26]	-0,084	0,034	0,005	-0,160	-0,085	-0,023
W[27]	-0,171	0,028	0,003	-0,224	-0,172	-0,108
W[28]	-0,062	0,029	0,002	-0,120	-0,062	-0,006
W[29]	0,031	0,031	0,002	-0,030	0,032	0,094
W[30]	0,137	0,034	0,003	0,071	0,138	0,200
W[31]	0,115	0,036	0,002	0,047	0,115	0,181
W[32]	0,158	0,027	0,004	0,105	0,156	0,208
W[33]	0,169	0,029	0,002	0,114	0,168	0,225
W[34]	0,127	0,030	0,004	0,075	0,127	0,182
W[35]	0,116	0,033	0,004	0,058	0,115	0,190
W[36]	0,120	0,034	0,003	0,047	0,122	0,185
W[37]	0,075	0,024	0,002	0,031	0,076	0,118
W[38]	-0,019	0,038	0,003	-0,098	-0,021	0,054



SEMINAR HASIL TESIS

Model *Mixture Survival* Spasial Pada Angka Lama Sekolah Anak Umur 16-18 Tahun di Jawa Timur Tahun 2012

Oleh :

Jihadil Qudsi S. (1313 201 033)

Pembimbing : Prof. Drs. Nur Iriawan, M.Ikom., Ph.D.

**PROGRAM MAGISTER
JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
2015**



Pendahuluan

Tinjauan Pustaka

Metodologi Penelitian

Hasil & Pembahasan

Penutup



PENDAHULUAN

Latar Belakang

Rumusan Masalah

Tujuan Penelitian

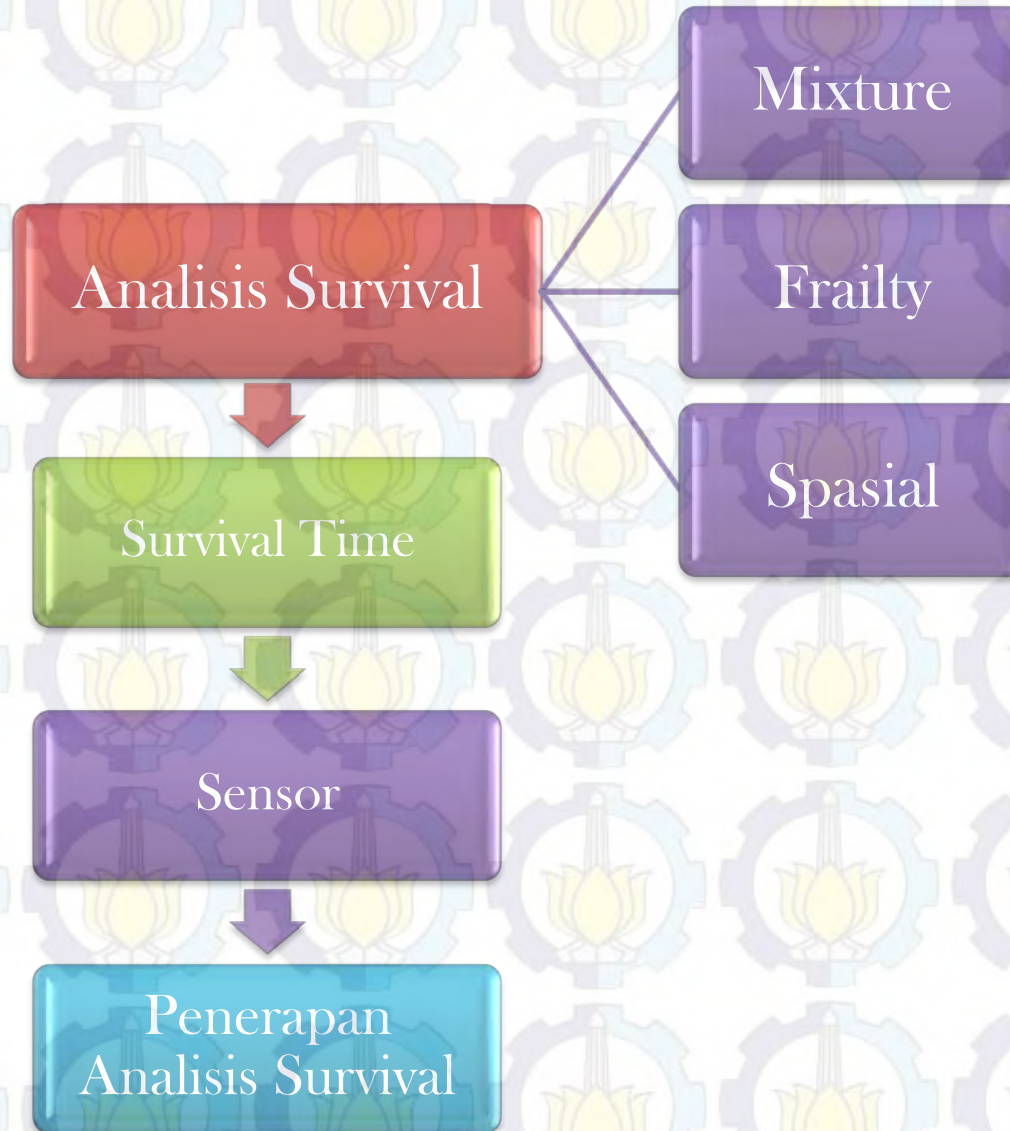
Manfaat Penelitian

Batasan Penelitian





LATAR BELAKANG





LATAR BELAKANG

Penerapan
Analisis
Survival

Pendidikan

IPM

WAJAR
9 Tahun

RLS

Angka
Lama
Sekolah

APS Jawa Timur tahun 2009-2012

2009			2010			2011			2012		
7-12	13-15	16-18	7-12	13-15	16-18	7-12	13-15	16-18	7-12	13-15	16-18
98,57	88,00	58,44	98,74	88,82	59,39	98,27	90,11	58,54	98,65	91,62	61,87



PENELITIAN SEBELUMNYA

Mixture Survival

- (Ando, T., Imoto, S., & Miyano, S., 2004)
- (Muthen, B., & Masyn, K. 2005)
- (Zhang 2008)
- (Hariyanto 2009)

Survival Spasial Frailty

- (Carlin dan Louis 2000)
- (Li dan Ryan 2002)

Mixture Survival Spasial Frailty

- (Berry dan Starr (1990 dan 1991))
- (Banerjee, Wall, & Carlin, 2003)
- (Hasyim 2012)

Angka Lama Sekolah

- (Santoso 2009) : Pendekatan Spline Multivariable Dan MARS.
- (Sulistiyawati 2009). Model Mixture Survival



RUMUSAN MASALAH & TUJUAN PENELITIAN

Rumusan Masalah

Bagaimana model *survival* dengan penambahan konsep *mixture* dan pertimbangan adanya efek spasial menggunakan pendekatan bayesian?

Bagaimana model *mixture survival* spasial pada angka lama sekolah anak umur 16-18 tahun di Jawa Timur berdasarkan variabel-variabel yang mempengaruhi tinggi atau rendahnya angka lama sekolah?

Tujuan Penelitian

Memperoleh model *survival* dengan penambahan konsep *mixture* dan pertimbangan adanya efek spasial menggunakan pendekatan bayesian.

Memperoleh model *mixture survival* spasial pada angka lama sekolah anak umur 16-18 tahun di Jawa Timur berdasarkan variabel-variabel yang mempengaruhi tinggi atau rendahnya angka lama sekolah.



TINJAUAN PUSTAKA

• Analisis *Survival*

• Fungsi *Survival* dan Fungsi Hazard

• Model Hazard Proporsional

• Model *Survival* Spasial

• Model Lattice

• Penimbang Spasial

• Model *Mixture*

• Analisis Bayesian

• Likelihood dan Prior

• Markov Chain Monte Carlo

• RJMCMC

• Odds Ratio

• Angka Lama Sekolah



ANALISIS SURVIVAL

Tiga faktor dalam menentukan waktu survival (Zang, 2008):

Titik Awal

Waktu Akhir

Skala Pengukuran Waktu

Alasan data disebut tersensor (Kleinbaum dan Klein, 2005):

Lost of follow up

Drop Out

Termination of the study

Tiga tipe penyensoran (Collet, 2003)

Sensor kiri

Sensor kanan

Sensor interval



FUNGSI SURVIVAL DAN FUNGSI HAZARD



Fungsi *survival* ($S(t)$) dinyatakan sebagai peluang seorang individu untuk bertahan lebih lama dari suatu waktu t .

Fungsi *hazard* dijelaskan sebagai reaksi sesaat atau laju kegagalan (*failure*) sesaat ketika mengalami suatu *event* pada waktu ke- t .

$$S(t) = P(T \geq t) = 1 - F(t)$$

$$h(t) = \lim_{\Delta t \rightarrow 0} \left\{ \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \right\}$$

Hubungan antara fungsi survival dan fungsi hazard

$$H(t) = -\log S(t)$$



MODEL HAZARD PROPORSIONAL

Persamaan model *hazard* proporsional :

$$h(t) = h_0(t) \exp(\beta^T \mathbf{x}) \longrightarrow \text{Regresi Cox}$$

dimana, $\mathbf{x} = (x_1, x_2, \dots, x_p)$ adalah variabel prediktor pada model *hazard* proporsional

Regresi Cox

Tidak harus memenuhi distribusi apapun (Hosmer dan Lemenshow, 1999)

Asumsi : fungsi hazard harus selalu proporsional



MODEL SURVIVAL SPASIAL

Model Semiparametrik Cox dengan Hazard rate

$$h(t_{ij}; \mathbf{x}_{ij}\boldsymbol{\beta}) \propto h_0(t_{ij}) \exp(\mathbf{x}_{ij}^T \boldsymbol{\beta})$$

dimana :

t_{ij} : waktu hingga terjadinya suatu *event*.

\mathbf{x}_{ij} : vektor dari *covariate* yang mempengaruhi kejadian tersebut.

Penyertaan *frailty*/efek random

$$h(t_{ij}; \mathbf{x}_{ij}\boldsymbol{\beta}) \propto h_0(t_{ij}) \exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + W_i)$$



MODEL LATTICE

(Banerjee dan Carlin, 2003)

Model Lattice menggunakan informasi jarak dari ketetanggaan (*neighborhood*) atau kedekatan antara satu *region* dengan *region* yang lain serta menggabungkan informasi tentang wilayah-wilayah yang saling berdekatan/ber-tetangga dibandingkan dengan informasi jarak metrik.

(Wall, 2004)

Distribusi dari efek random \mathbf{W}^* didefinisikan sebagai

$$\mathbf{W}^* | \lambda \sim \text{CAR}(\lambda)$$

merupakan model *conditionally autoregressive* yang menyatakan adanya kebergantungan spasial pada susunan kovarian.



MATRIKS PENIMBANG SPASIAL

Beberapa metode yang mendefinisikan hubungan kebersinggungan (*contiguity*) antar daerah (LeSage, 1999)

- *Linear contiguity* (persinggungan tepi)
- *Rook contiguity* (persinggungan sisi)
- *Bishop contiguity* (persinggungan sudut)
- *Double linear contiguity* (persinggungan dua tepi)
- *Double rook contiguity* (persinggungan dua sisi)
- *Queen contiguity* (persinggungan sisi-sudut)



MODEL MIXTURE

-(McLachlan dan Basford, 1988) : Model *mixture* akan tampak dari data yang diamati dan data yang ada biasanya terdiri dari beberapa subpopulasi atau grup.

$$f(x | \boldsymbol{\pi}) = \pi_1 f_1(x) + \pi_2 f_2(x) + \cdots + \pi_k f_k(x)$$

dengan π_1 : Nilai proporsi dari komponen *mixture* pertama

π_2 : Nilai proporsi dari komponen *mixture* kedua

π_k : Nilai proporsi dari komponen *mixture* ke- k

sehingga $\pi_1 + \pi_2 + \cdots + \pi_k = 1$



ANALISIS BAYESIAN

Metode Bayesian ditemukan oleh seorang ilmuwan bernama **Thomas Bayes** (1702-1761), yang menemukan perlakuan matematika untuk masalah non trivial dari inferensi Bayesian, yang digunakan untuk penyelesaian kasus khusus yang kemudian dikenal dengan nama Teorema Bayesian.

Hubungan distribusi *posterior* dengan distribusi *prior* dan *likelihood* dapat dituliskan sebagai berikut :

Distribusi *posterior* \propto *likelihood* \times Distribusi *prior*

$$f(\theta | y) \propto f(y | \theta) f(\theta)$$



LIKELIHOOD DAN PRIOR MODEL MIXTURE

Jika terdapat data pengamatan sebanyak n yang terdekomposisi ke dalam k kelompok data yang masing-masing mempunyai distribusi maka fungsi *likelihood* model *mixture* tersusun dari beberapa *likelihood* data di setiap subpopulasi menurut distribusi masing-masing, seperti berikut :

$$l_{mix} = \prod_{i=1}^n f_{mix}(x_i | \theta)$$

$$l_{mix} = \prod_{i_1=1}^{n_1} \lambda_1 f(x_{i_1} | \theta_1) + \prod_{i_2=1}^{n_2} \lambda_2 f(x_{i_2} | \theta_2) + \dots + \prod_{i_k=1}^{n_k} \lambda_k f(x_{i_k} | \theta_k)$$

Proper prior atau improper prior

Informative prior atau non informative prior

Conjugate prior atau non conjugate prior

Pseudo prior



MARKOV CHAIN MONTE CARLO

Carlin (1992) : Pendekatan MCMC sangat efektif untuk mengurangi beban komputasi dalam menyelesaikan persamaan integrasi yang kompleks dan memungkinkan proses simulasi dengan mengambil sampel acak dari model stokastik yang sangat rumit.

Ntzoufras, (2009) : Ide dasar dari MCMC adalah membangkitkan data sampel dari distribusi posterior sesuai proses markov chain dengan menggunakan simulasi monte carlo secara iteratif sehingga diperoleh kondisi yang konvergen terhadap posterior.

Gibbs Sampling

Metropolis Hasting

Reversible Jump MCMC



REVERSIBLE JUMP MCMC

Metropolis Hasting

Green (1995)

Reversible Jump MCMC

Richardson dan Green (1997) : memanfaatkan algoritma RJMCMC untuk mendapatkan model *mixture* yang tepat bagi data yang terindikasi bersifat multimodal dengan banyak komponen penyusun *mixture* yang tidak diketahui.

Enam tipe pergerakan proses RJMCMC :

1. *updating* ω
2. *updating* θ
3. *updating* z
4. *updating* hyperparameter β
5. *split/merge* komponen-komponen
6. *birth/death* dari komponen kosong



ODDS RATIO

Hosmer dan Lemeshow (1999)

Odds ratio merupakan suatu ukuran yang digunakan untuk mengetahui tingkat resiko (kecenderungan) yaitu perbandingan antara *Odd* individu dengan kondisi variabel prediktor X pada kategori sukses dengan kategori gagal.

$$OR = \frac{h(t | x = 1)}{h(t | x = 0)} = \frac{h_0(t)e^{\beta}}{h_0(t)} = e^{\beta}$$

Tingkat kecepatan terjadinya *failure event* pada individu dengan kategori $x = 1$ sebesar e^{β} kali tingkat kecepatan terjadinya resiko peristiwa *failure event* pada individu dengan kategori $x = 0$.

Sedangkan untuk variabel kontinu, nilai dari e^{β} mempunyai interpretasi sebagai perbandingan *odds ratio* antara individu dengan nilai x lebih besar 1 satuan dibandingkan individu yang lainnya.



ANGKA LAMA SEKOLAH

Angka lama sekolah yaitu jumlah tahun belajar penduduk usia 15 tahun ke atas yang telah diselesaikan dalam pendidikan formal (tidak termasuk tahun yang mengulang) (indikator dalam sistem rujukan BPS).

$$ALS = \text{Tahun Konversi} + (\text{Kelas Tertinggi} - 1)$$



METODOLOGI PENELITIAN



Sumber Data

**Variabel
Penelitian**

**Metode
Penelitian**



SUMBER DATA

• Data yang digunakan dalam penelitian ini data sekunder dari Badan Pusat Statistik (BPS), Jawa Timur

Jenis Data

Data

- Data angka lama sekolah anak yang berumur 16-18 tahun yang didapat dari perhitungan berdasarkan hasil SUSENAS tahun 2012 Jawa Timur.

- Responden dinyatakan *tidak tersensor* apabila lulus jenjang pendidikan sekolah menengah pertama (SMP) atau yang sederajat (MTs, SMPLB dan paket B) baik negeri maupun swasta dan mendapatkan ijazah pada rentang umur 16-18 tahun.

Kriteria Data



VARIABEL PENELITIAN

Y

- Angka lama sekolah anak umur 16-18 tahun

X1

- Jenis Kelamin Responden

X2

- Status Perkawinan Responden

X3

- Status Bekerja Responden

X4

- Tingkat Pendidikan Kepala Rumah Tangga

X5

- Jenis Kelamin Kepala Rumah Tangga

X6

- Rata-Rata Pengeluaran Rumah Tangga per Kapita per Bulan

X7

- Jumlah Anggota Rumah Tangga

X8

- Lokasi Tempat Tinggal Responden



LANGKAH-LANGKAH PENELITIAN

1. Mengkaji model *mixture survival* dengan mempertimbangkan adanya pengaruh lokasi (*spasial*) menggunakan pendekatan bayesian :

2. Menentukan model *mixture survival* spasial pada angka lama sekolah anak umur 16-18 tahun di Jawa Timur berdasarkan faktor-faktor yang mempengaruhi tinggi atau rendahnya angka lama sekolah anak umur 16-18 tahun :





HASIL DAN PEMBAHASAN

Model *Survival* Dengan Penambahan Konsep *Mixture* dan Pertimbangan Adanya Efek Spasial Menggunakan Pendekatan Bayesian

Model *Mixture Survival* Spasial Pada Angka Lama Sekolah Anak Umur 16-18 Tahun di Jawa Timur Berdasarkan Variabel-Variabel Yang Mempengaruhi.



MODEL *SURVIVAL* DENGAN PENAMBAHAN KONSEP *MIXTURE* DAN PERTIMBANGAN ADANYA EFEK SPASIAL MENGGUNAKAN PENDEKATAN BAYESIAN

Pendekatan lattice digunakan untuk menyusun susunan spasial dari daerah-daerah. Sehingga W_j terganti oleh W_j^* , hal tersebut dikarenakan \mathbf{W}^* didefinisikan oleh daerah-daerah dengan indeks diskret dimana daerah-daerah tersebut merupakan partisi/bagian dari daerah.

$$h(t_{ij}; \mathbf{x}\boldsymbol{\beta}) \propto h_0(t_{ij}) \exp(\mathbf{x}_{ij}^T \boldsymbol{\beta})$$



$$h(t_{ij}; \mathbf{x}\boldsymbol{\beta}) \propto h_0(t_{ij}) \exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + W_j^*)$$

dimana : $\mathbf{W}^* | \lambda \sim \text{CAR}(\lambda)$, $\lambda = 1/v$.



MODEL *SURVIVAL* DENGAN PENAMBAHAN KONSEP *MIXTURE* DAN PERTIMBANGAN ADANYA EFEK SPASIAL MENGGUNAKAN PENDEKATAN BAYESIAN

Pendekatan untuk setiap komponen *mixture* dari angka lama sekolah anak umur 16-18 tahun dilakukan dengan menggunakan distribusi Normal 2-parameter (μ, σ) sehingga,

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right)$$

$$F(t) = \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{t-\mu}{\sqrt{2}\sigma}\right)$$

$$S(t) = \frac{1}{2} - \frac{1}{2} \operatorname{erf}\left(\frac{t-\mu}{\sqrt{2}\sigma}\right)$$

$$h(t) = \frac{\sqrt{2} \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right)}{\sqrt{\pi\sigma^2} \left(1 - \operatorname{erf}\left(\frac{t-\mu}{\sqrt{2}\sigma}\right)\right)}$$

$$h(t, \mathbf{x}) = h_0(t) \exp(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + W_j^*)$$

$$= \frac{\sqrt{2} \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right)}{\sqrt{\pi\sigma^2} \left(1 - \operatorname{erf}\left(\frac{t-\mu}{\sqrt{2}\sigma}\right)\right)}$$



MODEL *SURVIVAL* DENGAN PENAMBAHAN KONSEP *MIXTURE* DAN PERTIMBANGAN ADANYA EFEK SPASIAL MENGGUNAKAN PENDEKATAN BAYESIAN

Penentuan distribusi prior sebagai berikut,

$$\sigma \sim \text{Gamma}(p, q)$$

$$\beta_i \sim \text{Normal}(m, s)$$

penentuan distribusi prior tersebut dilakukan berdasarkan gabungan antara *prior conjugate* dan *prior informatif*.

Penentuan join distribusi posterior

t -> waktu *survival* responden (waktu yang diperlukan hingga terjadinya lulus/berhenti sekolah),

γ_{ij} -> menyatakan status tersensor dari responden angka lama sekolah yaitu nilai 0 untuk responden tersensor dan nilai 1 untuk responden tidak tersensor.

x -> merupakan vektor *covariate* (variabel prediktor).

$$P(\beta, W, \sigma, \lambda \mid t, x, \gamma) \propto L(\beta, W, \sigma; t, x, \gamma) p(W \mid \lambda) p(\beta) p(\sigma) p(\lambda)$$



MODEL *MIXTURE SURVIVAL* SPASIAL PADA ANGKA LAMA SEKOLAH ANAK UMUR 16-18 TAHUN DI JAWA TIMUR BERDASARKAN VARIABEL-VARIABEL YANG MEMPENGARUHI

Karakteristik responden anak umur 16-18 tahun berdasarkan variabel-variabel prediktor

Variabel	Minimum	Maksimum	Mean	Std. Deviation
Angka Lama Sekolah Anak Umur 16-18 Tahun	0	12	9,39	1,84
Jumlah Anggota Rumah Tangga	1	11	4,41	1,34
Rata-Rata Pengeluaran Rumah Tangga	1,21	131,13	5,17	4,81



MODEL *MIXTURE SURVIVAL* SPASIAL PADA ANGKA LAMA SEKOLAH ANAK UMUR 16-18 TAHUN DI JAWA TIMUR BERDASARKAN VARIABEL-VARIABEL YANG MEMPENGARUHI

Tingkat Pendidikan Terakhir

Status Ber

Bekerja
20%

Lokasi Rumah Responden

Desa
44%

Kota
56%

Kepala Rumah
tangga

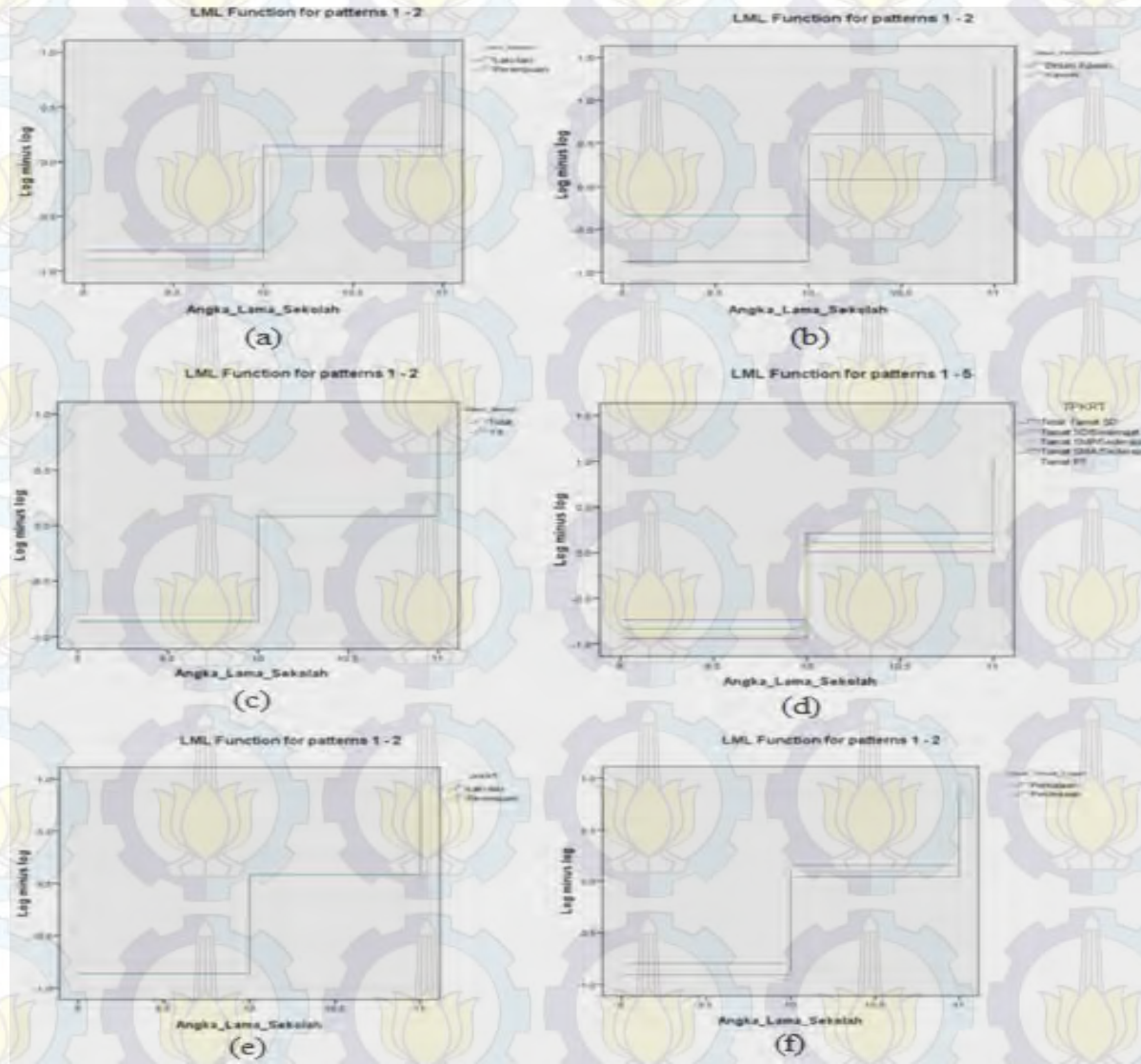
Laki-Laki
90%

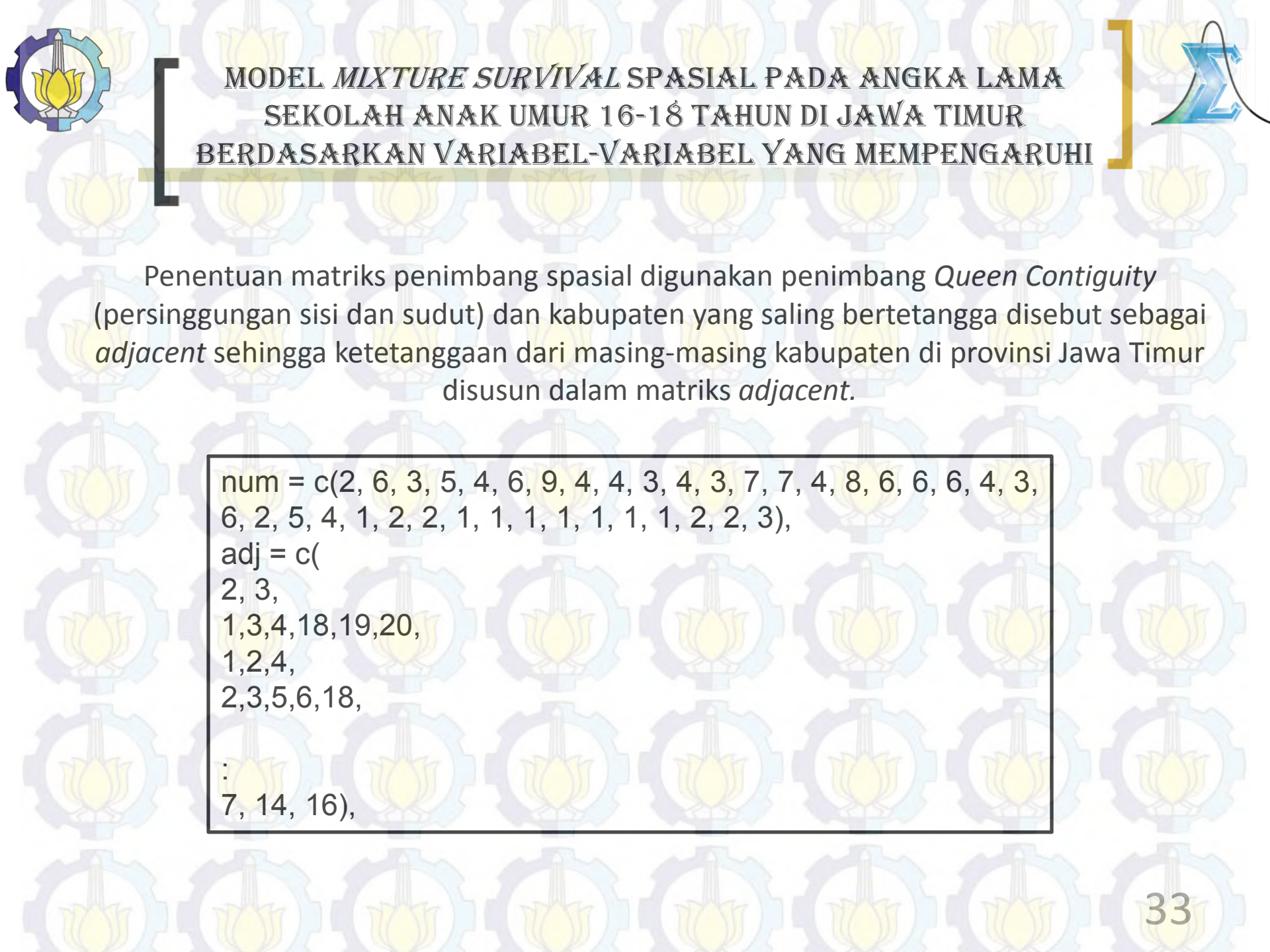
0

	0	1	2	3	4	5	6	7	8	9	10	11	12
Laki-laki	16	5	7	8	7	6	152	69	134	597	505	358	136
Perempuan	7	3	2	7	4	4	130	32	125	462	476	413	129



MODEL *MIXTURE SURVIVAL* SPASIAL PADA ANGKA LAMA SEKOLAH ANAK UMUR 16-18 TAHUN DI JAWA TIMUR BERDASARKAN VARIABEL-VARIABEL YANG MEMPENGARUHI





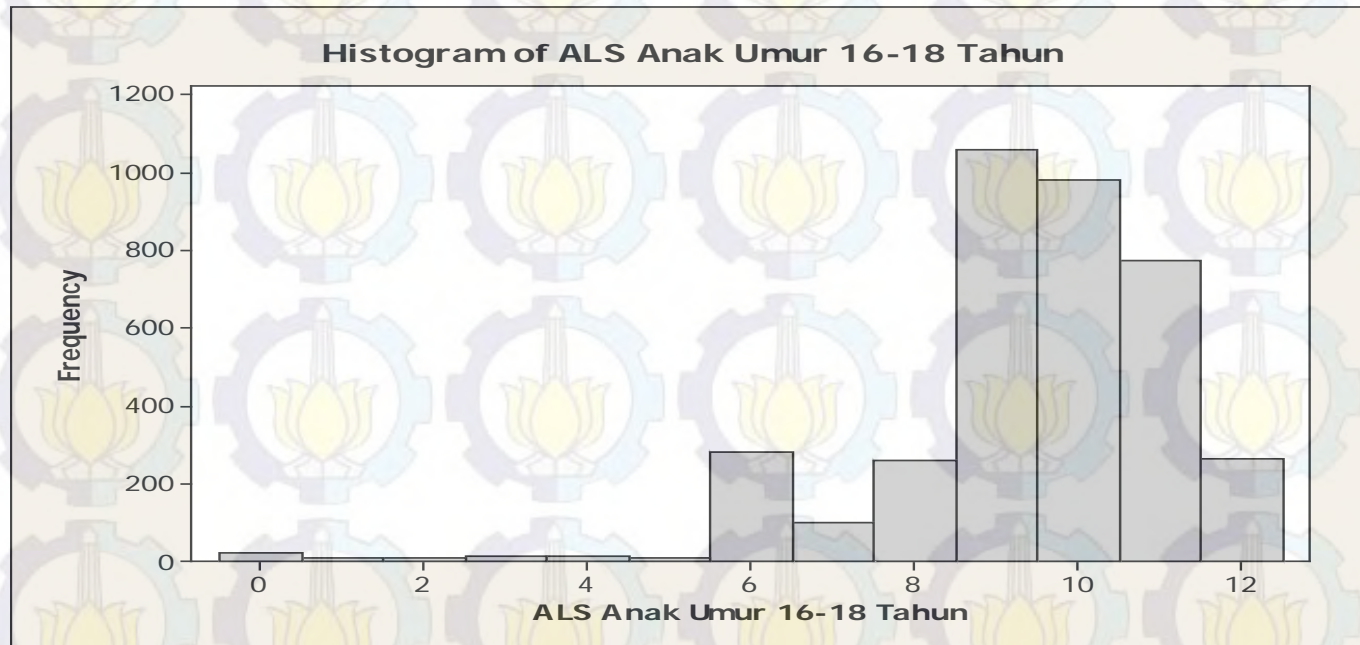
MODEL *MIXTURE SURVIVAL* SPASIAL PADA ANGKA LAMA SEKOLAH ANAK UMUR 16-18 TAHUN DI JAWA TIMUR BERDASARKAN VARIABEL-VARIABEL YANG MEMPENGARUHI

Penentuan matriks penimbang spasial digunakan penimbang *Queen Contiguity* (persinggungan sisi dan sudut) dan kabupaten yang saling bertetangga disebut sebagai *adjacent* sehingga ketetanggaan dari masing-masing kabupaten di provinsi Jawa Timur disusun dalam matriks *adjacent*.

```
num = c(2, 6, 3, 5, 4, 6, 9, 4, 4, 3, 4, 3, 7, 7, 4, 8, 6, 6, 6, 4, 3,  
6, 2, 5, 4, 1, 2, 2, 1, 1, 1, 1, 1, 1, 1, 2, 2, 3),  
adj = c(  
2, 3,  
1,3,4,18,19,20,  
1,2,4,  
2,3,5,6,18,  
:  
7, 14, 16),
```




MODEL *MIXTURE SURVIVAL* SPASIAL PADA ANGKA LAMA SEKOLAH ANAK UMUR 16-18 TAHUN DI JAWA TIMUR BERDASARKAN VARIABEL-VARIABEL YANG MEMPENGARUHI



Distribusi	Statistik Uji (A_n^2)	Nilai Kritis ($a_{n,1-\alpha}$)	Keputusan
Log Normal	281,78	2,5018	Tolak H_0
Normal	146,06	2,5018	Tolak H_0
Weibull	147,00	2,5018	Tolak H_0
Log Logistik	193,50	2,5018	Tolak H_0



MODEL MIXTURE SURVIVAL SPASIAL PADA ANGKA LAMA SEKOLAH ANAK UMUR 16-18 TAHUN DI JAWA TIMUR BERDASARKAN VARIABEL-VARIABEL YANG MEMPENGARUHI

```
model {  
  for (i in 1:n) {  
    T[i] ~ dnorm(psi[i], tau)I(t.cen[i],)  
  }  
  psi[1:n] <- jump.lin.pred(X[1:n, 1:Q], k,  
    beta.prec)  
  id <- jump.model.id(psi[1:n])  
  beta.prec <- tau / lambda  
  tau ~ dgamma(a, b)  
  k ~ dbin(0.5, Q)}  
list(  
  n = 3794, Q = 11,  
  lambda = 10000,  
  a = 0.001, b = 0.001,  
  T = c(0,0,...,12,12), t.cen=c(0,0,...,1,1),  
  X = structure(  
    .Data = c(  
      2,1,1,1,0,0,0,1,1,1.89602,2,  
      1,1,2,1,0,0,0,1,6,2.51399,2,  
      :  
      :  
      1,1,1,0,1,0,0,1,5,4.97648,1,  
      1,1,2,0,1,0,0,1,4,3.19188,1  
    ),  
    .Dim = c(3794, 11)))  
list(tau = 1)
```

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
id	1335.0	123.2	18.22	1310.0	1311.0	1823.0	1	100000
k	6.012	0.4016	0.06022	6.0	7.0	8.0	1	100000
tau	0.3387	0.01649	7.269E-4	0.3251	0.3393	0.3541	1	100000

model structure	posterior prob.	cumulative prob.
11111000101	0.86021	0.86021
11111000111	0.05878	0.91899
11110000111	0.05303	0.97202
11110000101	0.01396	0.98598
01111000101	0.00557	0.99155
variable no.	marginal prob.	
1	0.99296	
2	0.99998	
3	0.99989	
4	1.0	
5	0.92995	
6	6.0E-4	
7	9.1E-4	
8	1.7E-4	
9	0.99616	
10	0.11668	
11	0.99826	



```
model;
{
  for( i in 1 : N ) {T[i] ~ dnorm(psi[i],tau[i])|(t.cen[i],)}
  for( i in 1 : N ) {P[i] ~ dcat(Phi[1:6])}
  for( i in 1 : N ) {tau[i] <- ptau[P[i]]}
  for( i in 1 : N ) {psi[i] <- exp(b1[P[i]]*X1[i] + b2[P[i]]*X2[i] + b3[P[i]]*X3[i] + b4_1[P[i]]*X4_1[i] + b4_2[P[i]]*X4_2[i] +
b4_3[P[i]]*X4_3[i] + b4_4[P[i]]*X4_4[i] + b5[P[i]]*X5[i] + b6[P[i]]*X6[i] + b7[P[i]]*X7[i] + b8[P[i]]*X8[i] + W[kab[i]]))}
  for( i in 1:sumNumNeigh) {weights[i] <- 1}
  W[1:s] ~ car.normal(adj[], weights[], num[], dtau)
Phi[1:6] ~ ddirch(Alpha[])
ptau[1] ~ dgamma(0.1,0.1)
:
ptau[6] ~ dgamma(0.1,0.1)
psigma[1] <- 1 / sqrt(ptau[1])
:
psigma[6] <- 1 / sqrt(ptau[6])
b1[1] ~ dnorm(-0.167,0.4649)
:
b8[1] ~ dnorm(-0.296,0.4421)
:
:
b1[6] ~ dnorm(0.048,0.2036)
b2[6] ~ dnorm(-1.171,10.721)
:
b8[6] ~ dnorm(0.366,0.7329)
sigma <- 1 / sqrt(dtau)}
DATA
list(N=3794,s=38,
num = c(2, 6, 3, 5, 4, 6, 9, 4, 4, 3, 4, 3, 7, 7, 4, 8, 6, 6, 6, 4, 3, 6, 2, 5, 4, 1, 2, 2, 1, 1, 1, 1, 1, 1, 1, 2, 2, 3),
adj = c(
2, 3,
:
7, 14, 16),
sumNumNeigh=140,Alpha=c(1,1,1,1,1,1), dtau=1)
T[] t.cen[] kab[] X1[] X2[] ... X8[] P[]
NA 0      2      2      1      ...      2      1
:
12 0      38      1      1      ...      1      6
END
Inits
list(W=c(0,0,0,...,0),b1=c(0,0,0,0,0,0),...,b8=c(0,0,0,0,0,0))
```




MODEL MIXTURE SURVIVAL SPASIAL PADA ANGKA LAMA SEKOLAH ANAK UMUR 16-18 TAHUN DI JAWA TIMUR BERDASARKAN VARIABEL-VARIABEL YANG MEMPENGARUHI



Node	Mean	Sd	MC Error	2.5%	Median	97.5%
Phi[1]	0,0087	0,0014	0,0332	0,0059	0,0086	0,0118
Phi[2]	0,0077	0,0013	0,0279	0,0052	0,00758	0,0106
Phi[3]	0,0642	0,0040	0,0913	0,0566	0,0641	0,072
Phi[4]	0,0917	0,0046	0,0001	0,0829	0,0916	0,101
Phi[5]	0,7984	0,0065	0,133	0,7856	0,7986	0,811
Phi[6]	0,0293	0,0027	0,0561	0,0241	0,02919	0,035
b1[1]	0,0271	0,0893	0,0127	-0,1669	0,0249	0,1964
b2[1]	0,9876	0,1508	0,0221	0,7003	0,9863	1,2880
b3[1]	1,3280	0,1621	0,0242	1,0910	1,3200	1,5990
b4_1[1]	-0,0871	0,1801	0,0263	-0,5033	-0,0793	0,2023
b4_2[1]	0,0786	0,3273	0,0438	-0,6468	0,0727	0,7583
b4_3[1]	-0,1805	0,2108	0,0302	-0,5683	-0,1531	0,1717
b4_4[1]	-0,0875	0,3676	0,0499	-0,7417	-0,1116	0,7962
b5[1]	0,4103	0,0929	0,0131	0,2095	0,4080	0,5642
b6[1]	0,0037	0,0216	0,0029	-0,0382	0,8450	0,0578
b7[1]	-0,0256	0,0202	0,0028	-0,0748	-0,0232	0,0040
b8[1]	-0,8775	0,0850	0,0123	-1,0150	-0,8855	-0,6988
b1[2]	0,5267	0,0262	0,0039	0,4790	0,5311	0,5612
b2[2]	0,3795	0,0173	0,0026	0,3456	0,3817	0,4089
b3[2]	-1,0280	0,0147	0,0021	-1,0510	-1,0320	-0,9999
b4_1[2]	0,0243	0,0720	0,0107	-0,0926	0,0442	0,1283
b4_2[2]	0,6229	0,0595	0,0089	0,5324	0,6130	0,7132
b4_3[2]	-1,3290	0,2244	0,0256	-1,8250	-1,3200	-0,8784
b4_4[2]	-1,7660	0,2069	0,0258	-2,2410	-1,7680	-1,4300
b5[2]	-0,8042	0,0220	0,0032	-0,8425	-0,8030	-0,7690
b6[2]	0,1589	0,0057	0,8440	0,1485	0,1600	0,1685
b7[2]	0,7912	0,0025	0,3610	0,7861	0,7922	0,7945
b8[2]	0,5496	0,0095	0,0014	0,5362	0,5468	0,5681

Node	Mean	Sd	MC Error	2.5%	Median	97.5%
b1[3]	0,5389	0,0116	0,0017	0,5177	0,5380	0,5625
b2[3]	0,2462	0,0104	0,0015	0,2274	0,2447	0,2639
b3[3]	0,1697	0,0145	0,0021	0,1403	0,1742	0,1933
b4_1[3]	0,0103	0,0193	0,0028	-0,0166	0,0045	0,0466
b4_2[3]	0,2694	0,0204	0,0030	0,2300	0,2701	0,3010
b4_3[3]	-1,024	0,1288	0,0181	-1,2660	-1,0210	-0,7419
b4_4[3]	0,1204	0,3459	0,0514	-0,5799	0,2927	0,5987
b5[3]	-0,224	0,0167	0,0025	-0,2509	-0,2278	-0,1948
b6[3]	0,5691	0,0027	0,4010	0,5652	0,5688	0,5729
b7[3]	0,3746	0,0036	0,5390	0,3694	0,3743	0,3811
b8[3]	0,0199	0,0044	0,6320	0,0107	0,0202	0,0270
b1[4]	0,0972	0,0389	0,0057	0,0265	0,0956	0,1833
b2[4]	0,3460	0,0511	0,0075	0,2371	0,3584	0,4167
b3[4]	0,0562	0,0372	0,0054	0,0014	0,0483	0,1394
b4_1[4]	1,2940	0,1004	0,0146	1,1360	1,2720	1,4920
b4_2[4]	1,2620	0,1105	0,0160	1,0890	1,2320	1,5150
b4_3[4]	1,3750	0,1336	0,0191	1,1490	1,3830	1,6170
b4_4[4]	1,302	0,1066	0,0157	1,103	1,298	1,514
b5[4]	0,1496	0,0814	0,012	0,0139	0,1631	0,2691
b6[4]	0,0704	0,0024	0,35	0,0652	0,0711	0,0736
b7[4]	0,0232	0,012	0,0017	0,0018	0,0233	0,0453
b8[4]	0,0706	0,0331	0,0048	0,0107	0,0671	0,1427



Node	Mean	Sd	MC Error	2.5%	Median	97.5%
b1[5]	0,1418	0,0111	0,0014	0,1194	0,1418	0,1627
b2[5]	0,2348	0,0151	0,0019	0,203	0,237	0,2607
b3[5]	0,1695	0,0097	0,0011	0,1528	0,1687	0,1882
b4_1[5]	0,603	0,0316	0,0043	0,5353	0,602	0,6664
b4_2[5]	0,6197	0,0325	0,0045	0,5482	0,6193	0,691
b4_3[5]	0,67	0,0315	0,0042	0,6084	0,6691	0,7387
b4_4[5]	0,69	0,03	0,004	0,6328	0,69	0,7594
b5[5]	0,2619	0,0141	0,0018	0,2347	0,2604	0,2922
b6[5]	0,0133	0,0006	0,0393	0,0121	0,0133	0,0144
b7[5]	0,0719	0,0038	0,508	0,0647	0,0717	0,0799
b8[5]	0,2039	0,0107	0,0014	0,1821	0,2045	0,2215
b1[6]	0,1225	0,0281	0,0035	0,0583	0,124	0,1707
b2[6]	0,7286	0,0918	0,0131	0,5384	0,7199	0,912
b3[6]	0,2573	0,03	0,0035	0,2015	0,2576	0,311
b4_1[6]	-0,049	0,0818	0,0101	-0,2002	-0,0462	0,1035
b4_2[6]	0,153	0,0635	0,0079	0,0397	0,1538	0,2768
b4_3[6]	0,0977	0,064	0,0077	-0,0201	0,095	0,2263
b4_4[6]	0,0525	0,0554	0,0069	-0,0491	0,0492	0,1588
b5[6]	0,1959	0,0389	0,005	0,1236	0,1953	0,2728
b6[6]	0,0049	0,0027	0,271	-0,816	0,005	0,0098
b7[6]	0,054	0,0124	0,0016	0,0262	0,0547	0,0779
b8[6]	0,5175	0,062	0,0085	0,3849	0,5171	0,6355
ptau[1]	3,66	4,141	0,443	0,1324	2,147	15,8
ptau[2]	0,196	0,406	0,0486	0,001	0,04954	1,423
ptau[3]	0,012	0,0063	0,0007	0,004	0,01007	0,03
ptau[4]	0,27	0,04	0,003	0,198	0,2693	0,36
ptau[5]	0,14	0,004	0,161	0,1317	0,1394	0,15
ptau[6]	0,37	0,05	0,003	0,2781	0,3719	0,5
sigma	0,36	0,04	0,0013	0,2837	0,3522	0,45

Node	Mean	Sd	Mc Error	2.5%	Median	97.5%
W[1]	-0,272	0,045	0,006	-0,378	-0,267	-0,191
W[2]	-0,348	0,038	0,006	-0,413	-0,357	-0,285
W[3]	0,002	0,033	0,003	-0,060	0,002	0,067
W[4]	-0,076	0,023	0,003	-0,119	-0,076	-0,027
W[5]	-0,022	0,031	0,004	-0,079	-0,023	0,040
W[6]	-0,009	0,024	0,003	-0,057	-0,009	0,034
W[7]	-0,030	0,028	0,003	-0,078	-0,032	0,027
W[8]	0,058	0,031	0,003	0,987	0,058	0,121
W[9]	0,066	0,024	0,003	0,023	0,065	0,112
W[10]	-0,014	0,034	0,005	-0,070	-0,021	0,063
W[11]	0,046	0,034	0,004	-0,016	0,044	0,121
W[12]	0,086	0,039	0,005	0,017	0,086	0,165
W[13]	0,005	0,023	0,003	-0,042	0,005	0,052
W[14]	-0,055	0,037	0,005	-0,110	-0,069	0,010
W[15]	0,082	0,024	0,002	0,033	0,082	0,129
W[16]	-0,010	0,022	0,002	-0,052	-0,010	0,032
W[17]	0,091	0,024	0,002	0,042	0,090	0,137
W[18]	-0,043	0,025	0,004	-0,085	-0,045	-0,003
W[19]	0,011	0,027	0,002	-0,041	0,008	0,070
W[20]	0,041	0,038	0,002	-0,039	0,043	0,112
W[21]	0,039	0,030	0,003	-0,019	0,038	0,096
W[22]	-0,217	0,017	0,002	-0,249	-0,214	-0,192
W[23]	-0,014	0,026	0,002	-0,064	-0,013	0,039
W[24]	-0,171	0,028	0,004	-0,211	-0,178	-0,115
W[25]	0,046	0,023	0,002	0,001	0,046	0,091
W[26]	-0,084	0,034	0,005	-0,160	-0,085	-0,023
W[27]	-0,171	0,028	0,003	-0,224	-0,172	-0,108
W[28]	-0,062	0,029	0,002	-0,120	-0,062	-0,006
W[29]	0,031	0,031	0,002	-0,030	0,032	0,094
W[30]	0,137	0,034	0,003	0,071	0,138	0,200
W[31]	0,115	0,036	0,002	0,047	0,115	0,181
W[32]	0,158	0,027	0,004	0,105	0,156	0,208
W[33]	0,169	0,029	0,002	0,114	0,168	0,225
W[34]	0,127	0,030	0,004	0,075	0,127	0,182
W[35]	0,116	0,033	0,004	0,058	0,115	0,190
W[36]	0,120	0,034	0,003	0,047	0,122	0,185
W[37]	0,075	0,024	0,002	0,031	0,076	0,118
W[38]	-0,019	0,038	0,003	-0,098	-0,021	0,054



**MODEL *MIXTURE SURVIVAL* SPASIAL PADA ANGKA LAMA
SEKOLAH ANAK UMUR 16-18 TAHUN DI JAWA TIMUR
BERDASARKAN VARIABEL-VARIABEL YANG MEMPENGARUHI**

Node	Odds Ratio
<i>Mixture 4 (0,092)</i>	
b1[4]	1,10
b2[4]	1,41
b3[4]	1,06
b4_1[4]	3,65
b4_2[4]	3,53
b4_3[4]	3,96
b4_4[4]	3,68
b5[4]	1,16
b6[4]	1,07
b7[4]	1,02
b8[4]	1,07

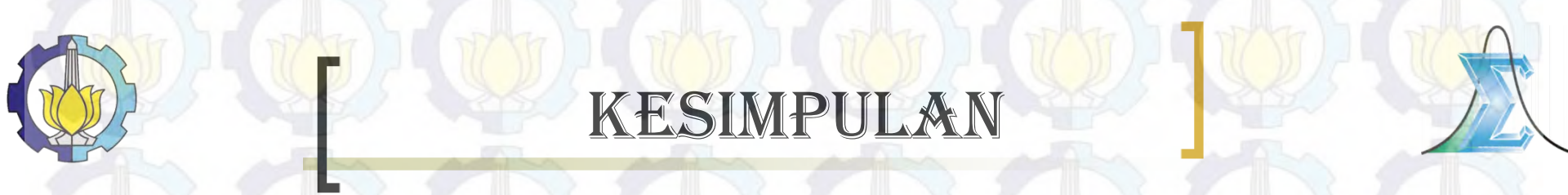


MODEL *MIXTURE SURVIVAL* SPASIAL PADA ANGKA LAMA SEKOLAH ANAK UMUR 16-18 TAHUN DI JAWA TIMUR BERDASARKAN VARIABEL-VARIABEL YANG MEMPENGARUHI

Jadi model *mixture survival* spasial yang terbentuk yaitu

$$\begin{aligned} h(t, \mathbf{x}) = & 0,087(h_1(t) \exp(0,98x_2 + 1,33x_3 + 0,41x_5 + 0,88x_8 + W_j^*)) \\ & + 0,008(h_2(t) \exp(0,53x_1 + 0,38x_2 - 1,03x_3 + 0,62x_{4_2} \\ & - 1,33x_{4_3} - 1,77x_{4_4} - 0,8x_5 + 0,16x_6 + 0,8x_7 + 0,55x_8 + W_j^*)) \\ & + 0,06(h_3(t) \exp(0,54x_1 + 0,25x_2 + 0,17x_3 + 0,27x_{4_2} - 1,02x_{4_3} \\ & - 0,22x_5 + 0,57x_6 + 0,37x_7 + 0,02x_8 + W_j^*)) \\ & + 0,092(h_4(t) \exp(0,1x_1 + 0,34x_2 + 0,06x_3 + 1,29x_{4_1} + 1,26x_{4_2} \\ & + 1,38x_{4_3} + 1,3x_{4_4} + 0,15x_5 + 0,07x_6 + 0,02x_7 + 0,07x_8 + W_j^*)) \\ & + 0,798(h_5(t) \exp(0,14x_1 + 0,23x_2 + 0,17x_3 + 0,6x_{4_1} + 0,62x_{4_2} \\ & + 0,67x_{4_3} + 0,69x_{4_4} + 0,26x_5 + 0,013x_6 + 0,072x_7 + 0,2x_8 + W_j^*)) \\ & + 0,029(h_6(t) \exp(0,122x_1 + 0,73x_2 + 0,26x_3 + 0,15x_{4_2} + 0,2x_5 \\ & + 0,05x_7 + 0,52x_8 + W_j^*)) \end{aligned}$$

dimana $h_1(t)$ hingga $h_6(t)$ adalah fungsi baseline hazard dari setiap komponen mixture dan W_j^* adalah efek spasial dari daerah ke- j .



KESIMPULAN

1. Model umum *mixture survival* spasial dimana pendekatan setiap komponen *mixture* dengan distribusi normal didapatkan 6 komponen *mixture* dan efek spasial W_j^* didapatkan sebagai berikut,

$$\begin{aligned} h(t, \mathbf{x}) = & \pi_1(h_1(t) \exp(\beta_1^T x_{p1} + W_j^*)) + \pi_2(h_2(t) \exp(\beta_2^T x_{p2} + W_j^*)) \\ & + \pi_3(h_3(t) \exp(\beta_3^T x_{p3} + W_j^*)) + \pi_4(h_4(t) \exp(\beta_4^T x_{p4} + W_j^*)) \\ & + \pi_5(h_5(t) \exp(\beta_5^T x_{p5} + W_j^*)) + \pi_6(h_6(t) \exp(\beta_6^T x_{p6} + W_j^*)) \end{aligned}$$

2. Variabel-variabel yang secara signifikan mempengaruhi angka lama sekolah anak umur 16-18 tahun berbeda pada 6 komponen mixture. Serta terdapat 23 kabupaten/kota memiliki efek spasial yang secara signifikan mempengaruhi angka lama sekolah anak umur 16-18 tahun.



Daftar Pustaka

- Aksioma, D. F. (2011). *Model Spasial Survival dengan Pendekatan Bayesian (Studi Kasus pada Kejadian HIV/AIDS di Provinsi Jawa Timur)*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Aksioma, D. F., & Iriawan, N. (2010). Spatial Autocorrelation of the DHF Outbreaks in the City of Surabaya. *Proceeding of The Third International Conference on Mathematics and Natural Science (ICMNS) 2010*. Bandung.
- Ando, T., Imoto, S., & Miyano, S. (2004). Kernel Mixture Survival Models for Identifying Cancer Subtypes, Predicting Patient's Cancer Types and Survival Probabilities. *Genome Informatics*, 15 (2), 201-210.
- Banerjee, S., Wall, M. M., & Carlin, B. P. (2003). Frailty Modeling for Spatially Correlated Survival Data, With Application to Infant Mortality in Minnesota. *Biostatistics*, 123-142.
- Besag, J., York, J., & Mollie, A. (1991). Bayesian Image Restoration with Two Application in Spatial Statistics. *Annals of the Institute of Statistical Mathematics*, 1-59.
- Best, N., Mason, A., dan Li, P. (2011), *Bayesian hierarchical modelling using WinBUGS*. Economic and Social Research Council, ESRC National Centre for Research Methods, UK.
- Boldstad, W.M., (2007). *Introduction to Bayesian Statistics. 2nd Edition*. New Jersey: Wiley.
- Box, G. E. & Tiao, G. C., 1973. *Bayesian Inference in Statistical Alaysis*. London: Addison-Wisley.
- BPS. (2014). *sirusa.bps.go.id*. Retrieved October 6, 2014, from sirusa.bps.go.id: <http://sirusa.bps.go.id>
- BPS. (2012). *Statistik Indonesia 2012*. Jakarta: BPS.
- BPS. (2012). *Pedoman Survei Sosial Ekonomi Nasional 2013*. Jakarta: BPS.
- Carlin, B. P., & Louis, T. A. (2000). *Bayes And Empirical Bayes Methods For Data Analysis (2ed)*. Boca Raton, FL: Chapman and Hall/CRC Press.
- Collet, D. (1994). *Modelling Survival Data in Medical Research*. London: Chapman and Hall.
- Collet, D. (2003). *Modelling Survival Data in Medical Research 2nd ed*. London: Chapman and Hall.
- Cox, D., & Oakes, D. (1984). *Analysis of Survival Data*. London: Chapman and Hall.
- Cressie, N. A. C. (1993). *Statistics for Spatial Data*, revised edition. New York: Wiley.
- Darmofal, D. (2008). *Bayesian Spatial Survival Models for Political Event Processes*. Columbia: Departement of Political Science, University of South Carolina 350 Gambrel Hall.
- Ducrocq, V. (1997). *Survival Analysis, A Statistical Tool For Longevity Data, 48th Annual Meeting Of The European Assosiation For Animal Production*. Vienna: Institut National de la Recherche Agronomique.



Daftar Pustaka

- Gamerman, D., 1997. *Markov Chain Monte Carlo*. London: Chapman & Hall.
- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995). *Bayesian Data Analysis* (2nd ed.). London: Chapman and Hall.
- Gelman, A., & Pardoe, L. (2006), *Bayesian Measure of Explained Variance and Pooling in Multilevel (Hierarchical) Models*, *Technometrics*, **48**(2): 241-251.
- Hariyanto, S. (2009). *Model Mixture Survival pada Kasus Lama Mencari Kerja di Pulau Jawa Tahun 2007*, Tesis. Surabaya: Institut Teknologi Sepuluh Nopember.
- Hasyim, M. (2012). *Model Mixture Survival Spasial dengan Frailty Berdistribusi Conditionally Autoregressive (CAR) pada Kasus Kejadian Demam Berdarah Dengue (DBD) di Kabupaten Pamekasan*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Hobcraft, J., McDonald, J., Menken, J., Rodriguez, G., & Trussel, J. (1984). *A Comparative Analysis of determinants of Birth Intervals. In, WFS Comparative Study: Cross-National Summaries (World Fertility Surveys)*. Voorburg, Netherland: International Statistical Institute.
- Hosmer Jr., D., & Lemeshow, S., .. (1999). *Applied Survival Analysis Regression Modelling of Time to Event Data*. New York: John Wiley and Sons. Inc.
- Ikana, M. (2005). *Pengaruh Urutan Kelahiran Terhadap Kelangsungan Pendidikan Anak Perempuan Usia 7-15 Tahun di Indonesia (Analisis Data Susenas 2002-KOR)*, Tesis. Depok: Universitas Indonesia.
- Iriawan, N. (2000). Pemilihan Model Menggunakan Struktur Perkalian Distribusi. *KAPPA: Jurnal Sains* , 37-44.
- Iriawan, N. (2001). *Studi Tentang Bayesian Mixture Normal dengan Menggunakan Metode MCMC*. Surabaya: Lemlit ITS.
- Ismartini, P. (2013). *Pengembangan Model Linear Hirarki Dengan Pendekatan Bayesian Untuk Pemodelan Data Pengeluaran Per Kapita Rumah Tangga*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Kartika, Y. (2007). *Pola Penyebaran Spasial Demam Berdarah Dengue di Kota Bogor Tahun 2005. Tugas Akhir*. Bogor: Institut Pertanian Bogor.
- Kleinbaum, D. G., & Klein, M. (2005). *Survival Analysis: A Self Learning* (3rd ed.). New York: Springer.
- Law, A. M., & Kelton, D. W. (2000). *Simulation Modeling and Analysis (3rd ed)* (3rd ed.). New York: McGraw-Hill.
- Lee, E. (1992). *Statistical Models and Methods for Lifetime Data*. New York: John Wiley and Sons. Inc.
- Lesage, J. P. (1999). *The Theory and Practice of Spatial Econometrics*. Toledo: Department of Economics, University of Toledo.



Daftar Pustaka

- Li, Y., & Ryan, L. (2002). Modelling Spatial Survival Data Using Semiparametric Frailty Models. *Biometrics*, 58, 287-297.
- Marin, J. M., Mengersen, K. & Robert, C. P., 2001. Bayesian Modelling and Inference on Mixture of Distribution. *Handbook of Statistics*, 25(50).
- McLachlan, G., & Basford, K. E. (1988). *Mixture Models Inference and Application to Clustering*. New York: Marcel Dekker.
- McLachlan, G., & Peel, D. (2000). *Finite Mixture Models*. New York: John Wiley and Sons, Inc.
- Miller, R. (1998). *Survival Analysis*. New York: John Willey and Sons Inc.
- Muthen, B., & Masyn, K. (2005). Discrete-Time Mixture Survival Analysis. *Journal of Educational and Behavioral Statistics*, 30(1), 27-58.
- Ntzoufras, I. (2009). *Bayesian Modelling Using WinBUGS*. New York: John Willey and Sons, Inc.
- Richardson, S., & Green, P. J. (1997). On Bayesian Analysis of Mixture with an Unknown Number of Component. *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 59, No. 4, 731-792.
- Santoso, B. (2009). *Pendekatan Spline Multivariable Dan MARS Untuk Pemodelan Lama Sekolah Pada Penduduk Usia Sekolah di Provinsi Papua, Tesis*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P., and Linde, A. (2002), *Bayesian measures of model complexity and fit*. Journal of The Royal Statistical Society, 64(1): 583-639.
- Sulistiyawati, D. (2009). *Model Mixture Survival Pada Kasus Lama Sekolah di Kabupaten Boalemo Provinsi Gorontalo, Tesis*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Vittinghoff, E., Glidden, D. V., Shiboski, S. C., & McCulloch, C. E. (2005). *Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models*. New York: Springer.
- Wackerly, D.D., Mendenhall, W., dan Scheaffer, R.L. (1996), *Mathematical Statistics with Applications*. Duxbury Press, California.
- Wall, M. M. (2004). A Close Look at The Spatial Structure Implied by the CAR and SAR Models. *Journal of Statistical Planning and Inference*, 121, 311-324.
- Zang. (2008). *Survival Analysis*. California: Wadsworth.
- Zhang, Y. (2008). *Parametric Mixture Models in Survival Analysis with Application, Disertation*. Philadelphia: Temple University.

Terima Kasih

